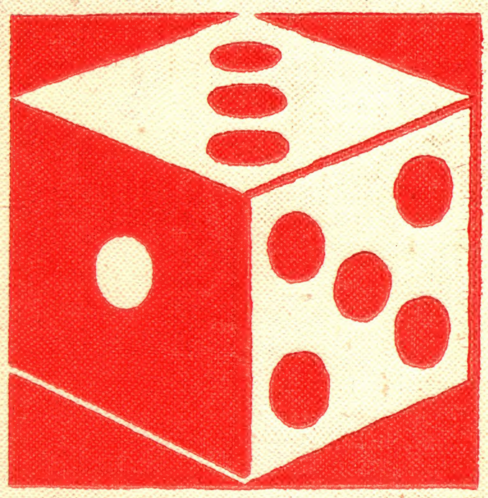
Л. Я. САВЕЛЬЕВ



КОМБИНАТОРИКА И ВЕРОЯТНОСТЬ

АКАДЕМИЯ НАУК СССР
СИБИРСКОЕ ОТДЕЛЕНИЕ
ИНСТИТУТ МАТЕМАТИКИ

Л. Я. САВЕЛЬЕВ

КОМБИНАТОРИКА И ВЕРОЯТНОСТЬ



ИЗДАТЕЛЬСТВО «НАУКА»
СИБИРСКОЕ ОТДЕЛЕНИЕ
Новосибирск · 1975

УДК 519.1, 519.2

Книга посвящена элементарной комбинаторике, теории вероятностей и их приложениям. В ней систематически используется теоретико-множественный язык. Абстрактность этого языка компенсируется большим количеством подробно разобранных примеров. Задачи собраны в отдельные части, которые можно читать независимо. Там рассматриваются простые модели, связанные с современными приложениями комбинаторики и теории вероятностей.

Книга предназначена для инженеров и научных работников всех специальностей, интересующихся вопросами теории вероятностей. Она будет полезна учащимся и преподавателям.

С $\frac{20204-712}{055(02)-75}$ 369—74

ПРЕДИСЛОВИЕ

Комбинаторика изучает способы подсчета числа элементов в различных множествах. К этому подсчету сводятся многие практические и теоретические задачи. Комбинаторные методы применяются в теории вероятностей, статистике, экономике, физике, химии, биологии и других науках. В книге описываются простейшие понятия комбинаторики и решаются некоторые связанные с ними классические задачи. Главной целью книги является изложение элементарной теории вероятностей.

Теорию вероятностей можно описательно определить как математическую теорию случайных явлений. С влиянием случая приходится сталкиваться при изучении самых различных явлений. Часто это влияние настолько существенно, что им нельзя пренебречь. Поэтому теория вероятностей используется почти во всех достаточно развитых областях науки, техники и производства. В книге излагается математическая часть элементарной теории вероятностей и дается представление о ее приложениях.

Математический язык позволяет избежать путаницы и ошибок при постановке и решении задач. Кроме того, он дает возможность формулировать определения и теоремы так, что они почти без изменений переносятся в общую теорию. Все это вполне компенсирует усилия, которые приходится затратить на изучение используемого математического языка.

Некоторое представление о приложениях теории вероятностей дают многочисленные подробно разобранные примеры и задачи. Примеры иллюстративного характера помещены в тех же главах, в которых рассматриваются соответствующие математические модели. Сравнительно более сложным и интересным задачам посвящена отдельная часть книги, в которой есть задачи, связанные с современными приложениями теории вероятностей к технике, экономике, биологии, медицине и психологии.

Книга «Комбинаторика и вероятность» состоит из двух разделов. Раздел «Комбинаторика» содержит две части. Первая посвящена элементам теории множеств и элементам комбинаторики. Вторая — задачам. Раздел «Вероятность» состоит из трех частей. Первая посвящена конечным вероятностным моделям. Вторая — случайным переменным. Третья — задачам. Если интересоваться исключительно приложениями, то можно читать только последние части каждого из разделов. В этих частях повторяются определения основных понятий и связанные с ними факты. Сравнительно сложные места книги отмечены звездочками. При первом чтении эти места рекомендуется пропустить.

Раздел I

КОМБИНАТОРИКА

Сколько элементов в данном множестве?

К этому вопросу сводится большое количество самых разных задач:

Сколько слагаемых в сумме, представляющей степень бинома?

Сколькими способами можно выбрать определенное количество из имеющихся предметов?

Сколькими способами могут распределиться частицы по областям пространства?

Сколько наборов может образоваться из данного двойного набора хромосом?

Комбинаторика изучает способы подсчета числа элементов в множестве. Она формулирует правила подсчета для некоторых простых множеств. Удачное использование этих правил позволяет подсчитывать число элементов в более сложных множествах.

Появление мощных вычислительных машин резко увеличило возможности комбинаторики и расширило ее приложения. Комбинаторные методы применяются сейчас в теории вероятностей, статистике, экономике, теории игр, физике, химии, биологии и многих других областях науки, техники и производства.

Часть I

ЭЛЕМЕНТЫ ТЕОРИИ МНОЖЕСТВ И КОМБИНАТОРИКИ

В этой части сначала кратко описывается язык теории множеств. Он очень удобен и широко используется в математике. Его точность окупает усилия, затрачиваемые на овладение им.

Затем излагаются некоторые простые правила подсчета числа элементов различных конечных множеств. С помощью этих правил решаются классические задачи о числе перестановок, сочетаний и размещений.

Глава 1

ЭЛЕМЕНТЫ ТЕОРИИ МНОЖЕСТВ

При первом чтении рекомендуется ограничиться параграфами 1—3 и 7 этой главы.

§ 1. МНОЖЕСТВА

Представление о понятии *много*, *множество* вырабатывается у каждого человека с раннего детства. Он постоянно сталкивается с различными множествами людей, животных или вещей: родственников, домашних животных, игрушек. В комбинаторике и теории вероятностей, как почти во всех разделах математики, используется так называемая наивная теория множеств. В этой теории понятие множества формально не определяется. Считается, что множество обладает всеми свойствами, которые ему приписываются. Уверенность в этом основывается на успешном применении математических методов в практической и теоретической деятельности.

1.1. Терминология

Каждое *множество* A определяется принадлежащими ему *элементами* a, b, c, ... Говорят также, что множество A образовано элементами a, b, c, ... или составлено из элементов a, b, c, ...

Пример 1. Множество L букв латинского алфавита определяется своими элементами a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z. Вместо «L является множеством букв латинского алфавита» пишут также

$$L = \{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z\}.$$

Пример 2. Множество F цифр определяется своими элементами 0, 1, 2, 4, 5, 6, 7, 8, 9. Вместо «F является множеством цифр» пишут также

$$F = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}.$$

Пример 3. Множество B исходов эксперимента, заключающегося в подбрасывании монеты один раз, определяется своими элементами 0 (монета падает цифрой вверх) и 1 (монета падает гербом вверх). Вместо «множество B образовано элементами 0, 1» пишут также

$$B = \{0, 1\}.$$

Пример 4. Множество U исходов эксперимента, заключающегося в подбрасывании игральной кости один раз, определяется своими элементами 1, 2, 3, 4, 5, 6, выражающими выпавшее число очков. Вместо «множество U образовано элементами 1, 2, 3, 4, 5, 6» пишут также

$$U = \{1, 2, 3, 4, 5, 6\}.$$

Пример 5. Множество решений квадратного уравнения

$$ax^2 + bx + c = 0.$$

Часто рассматривается *пустое* множество O, не имеющее элементов. Например, множество вещественных чисел, являющихся решениями квадратного уравнения

$$x^2 + 1 = 0,$$

— пустое множество. Чаще всего элементы обозначаются малыми, а множества — большими латинскими буквами.

Тот факт, что элемент a принадлежит множеству A, выражается символом

$$a \in A,$$

отрицание этого факта — символом

$$a \notin A.$$

В частности, для примера 1 выражение $a \in L$ означает, что буква a принадлежит латинскому алфавиту, а выражение $2 \notin L$,— что цифра 2 не принадлежит латинскому алфавиту.

Каждое множество, которому принадлежит ровно один элемент, называется *элементарным*. Например, множества $\{a\}$, $\{0\}$, $\{1\}$ являются элементарными. Множество решений квадратного уравнения $x^2-2x+1=0$ также является элементарным множеством.

2.1. Равенство множеств

Естественно считать равными множества, образованные одними и теми же элементами.

Определение равенства. *Множества A и B равны, если каждый элемент множества A принадлежит множеству B и каждый элемент множества B принадлежит множеству A.*

Например, множества $A=\{1, 2, 3, 4, 5, 6\}$ и $B=\{6, 5, 4, 3, 2, 1\}$ равны. Множество C решений квадратного уравнения

$$x^2-x=0$$

и множество чисел $D=\{0, 1\}$ тоже равны. Равенство множеств A и B записывается символом

$$A=B,$$

отрицание равенства множеств A и B — символом

$$A\neq B.$$

Отношение равенства для множеств, как и отношение равенства для чисел, обладает свойствами рефлексивности, симметричности и транзитивности. Это значит, что для каждых множеств A, B, C верны следующие предложения:

1) $A=A$ (*рефлексивность*);
2) если $A=B$, то $B=A$ (*симметричность*);
3) если $A=B$ и $B=C$, то $A=C$ (*транзитивность*).

§ 2. ОБЪЕДИНЕНИЕ И ПЕРЕСЕЧЕНИЕ МНОЖЕСТВ

Содержание понятий объединения и пересечения множеств ясно из названий этих операций. Объединением родителей и детей является семья. Пересечение русских и москвичей составляют русские, живущие в Москве.

1.2. Объединение множеств

Для каждых множеств A, B существует множество $A\cup B$, образованное всеми элементами, которые принадлежат хотя бы одному из множеств A или B.

Пример 1. Объединением множества $A=\{0, 2, 4, 6, 8\}$ четных цифр и множества $B=\{1, 3, 5, 7, 9\}$ нечетных цифр является множество $\cup B=\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ всех цифр.

Пример 2. Если $A = \{1, 2, 3\}$ и $B = \{2, 3, 4\}$, то $A \cup B = \{1, 2, 3, 4\}$.

Пример 3. Объединением множеств $A = \{1\}$ и $B = \{0, 1\}$ решений квадратных уравнений $x^2 - 2x + 1 = 0$ и $x^2 - x = 0$ является множество $A \cup B = \{0, 1\}$ решений уравнения $(x^2 - x)(x^2 - 2x + 1) = 0$.

Определение объединения. *Объединением множеств A, B называется множество $A \cup B$, образованное всеми элементами, которые принадлежат хотя бы одному из множеств A или B.*

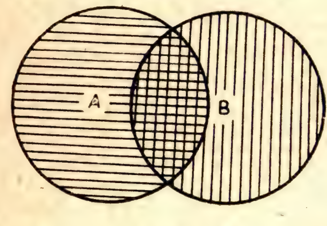


Рис. 1.

Определение объединения поясняется рис. 1. На этом рисунке левый круг с горизонтальной штриховкой изображает множество A, правый с вертикальной штриховкой — множество B, а вся заштрихованная фигура — объединение $A \cup B$ множеств A, B.

2.2. Пересечение множеств

Для каждых множеств A, B существует множество $A \cap B$, образованное всеми элементами, которые принадлежат каждому из множеств A и B.

Пример 1. Пересечением множества $A = \{0, 2, 4, 6, 8\}$ четных цифр и множества $B = \{1, 3, 5, 7, 9\}$ нечетных цифр является пустое множество $A \cap B = O$.

Пример 2. Если $A = \{1, 2, 3\}$ и $B = \{2, 3, 4\}$, то $A \cap B = \{2, 3\}$

Пример 3. Пересечением множеств $A = \{1\}$ и $B = \{0, 1\}$ решений квадратных уравнений $x^2 - 2x + 1 = 0$ и $x^2 - x = 0$ является множество $A \cap B = \{1\}$ решений системы квадратных уравнений $x^2 - 2x + 1 = 0$, $x^2 - x = 0$.

Определение пересечения. *Пересечением множеств A, B называется множество $A \cap B$, образованное всеми элементами, которые принадлежат каждому из множеств A и B.*

Пересечение $A \cap B$ множеств A и B обозначается также символом $A \cdot B$. Определение пересечения поясняется рис. 1: пересечение $A \cap B$ изображается дважды заштрихованной общей частью кругов A и B. Вместо *пересечение* говорят иногда *общая часть*.

Если пересечением множеств A и B является пустое множество O, то говорят, что множества A и B *не пересекаются*. В этом случае объединение $A \cup B$ множеств A и B называют также суммой и обозначают символом $A + B$.

§ 3. ЧАСТИ И ДОПОЛНЕНИЯ МНОЖЕСТВ

Содержание понятий части и дополнения множества ясно из названий этих операций. Например, совершеннолетние являются частью населения Советского Союза, а несовершеннолетние составляют дополнение этой части.

1.3. Части множества

Каждое множество A, образованное некоторыми элементами множества B, естественно назвать *частью* множества B.

Пример 1. Множество $A=\{0, 2, 4, 6, 8\}$ четных цифр является частью множества $B=\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ всех цифр.

Пример 2. Элементарное множество $A=\{1\}$ является частью множества $B=\{0, 1\}$.

Пример 3. Множество $A=\{1, 2\}$ является частью множества B решений уравнения $x^3-3x^2+2x=0$.

Определение части. *Произвольное множество A, каждый элемент которого принадлежит множеству B, называется частью множества B.*

Определение части поясняется рис. 2. Внутренний незаштрихованный круг обозначает множество A, являющееся частью множества B, которое изображается внешним кругом.

Рис. 2.

Вместо «множество A является *частью* множества B» часто говорят «A является *подмножеством* множества B». Говорят также «множество A *содержится* в множестве B» или «множество A включается в множество B» и пишут

$$A\subseteq B.$$

Отрицание того, что множество A содержится в множестве B, записывается символом

$$A\nsubseteq B.$$

Отношение «$\subseteq$» называют отношением *включения.*

По определению, пустое множество $A=O$ является частью каждого множества B. Множество $A=B$ является частью множества B. Пустое множество O и само множество B называются *несобственными* частями множества B. Остальные части множества B, не равные O или B, называются *собственными* частями множества B.

Если $A\subseteq B$ и $A\neq B$, то условимся говорить также, что «множество A *строго содержится* в множестве B», и писать $A\subset B$.

Пример. Частями множества $B=\{0, 1\}$ являются множества O, $\{0\}$, $\{1\}$, B. Собственными частями множества B являются элементарные множества $\{0\}$ и $\{1\}$. Таким образом: $O\subset B$, $\{0\}\subset B$, $\{1\}\subset B$, $B\subseteq B$.

2.3. Дополнение множества

Для каждой части A множества B существует часть $B-A$ множества B, образованная всеми элементами множества B, не принадлежащими множеству A.

Пример 1. Множество $B-A=\{1, 3, 5, 7, 9\}$ нечетных цифр является дополнением множества $A=\{0, 2, 4, 6, 8\}$ четных цифр до множества $B=\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ всех цифр.

Пример 2. Элементарное множество $B-A=\{0\}$ является дополнением элементарного множества $A=\{1\}$ до множества $B=\{0, 1\}$.

Пример 3. Множество $B-A=\{0\}$ является дополнением множества $A=\{1, 2\}$ до множества $B=\{0, 1, 2\}$ решений уравнения $x^3-3x^2+2x=0$.

Определение дополнения. *Дополнением части A множества B до множества B называется часть B—A множества B, образованная всеми элементами множества B, которые не принадлежат множеству A.*

Определение дополнения поясняется рис. 2. На этом рисунке заштрихованная часть внешнего круга изображает дополнение $B-A$ множества A до множества B.

В тех случаях, когда в тексте рассматриваются дополнения до одного и того же определенного множества U, условимся вместо «дополнение множества A до множества U» говорить коротко: «дополнение множества A» и вместо $U-A$ писать A'.

Пример. Если $U=\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ и $A=\{0, 2, 4, 6, 8\}$, то $A'=\{1, 3, 5, 7, 9\}$. Заметим, что

$$A\cap A'=\{0, 2, 4, 6, 8\}\cap\{1, 3, 5, 7, 9\}=0,$$

$$A\cup A'=\{0, 2, 4, 6, 8\}\cup\{1, 3, 5, 7, 9\}=\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}=U,$$

$$A''=\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}-\{1, 3, 5, 7, 9\}=\{0, 2, 4, 6, 8\}=A.$$

Равенства крайних множеств имеют общий характер: например, *дополнение к дополнению множества равно этому множеству.*

Рассмотрим множество U, его произвольную часть A и ее дополнение A'. Основные свойства дополнения выражает

Теорема о дополнении.

$$A\cap A'=0, \quad A\cup A'=U, \quad A''=A.$$

— Эти равенства вытекают непосредственно из определений пересечения объединения, дополнения и законов логики. Рис. 2 делает рассматриваемые равенства очевидными. Рассмотрим произвольный элемент x множества U.

1. Предположим, что пересечение множеств A и A' не пусто. По определению пересечения множеств из $x\in A\cap A'$ следует $x\in A$ и $x\in A'$, т. е. $x\in A$ и $x\notin A$, что невозможно. Следовательно, пересечение множеств A и A' пусто. Первое равенство доказано.

2. По определению дополнения $x\in A$ или $x\in A'$. По определению объединения это значит, что $x\in A\cup A'$. Таким образом, каждый элемент множества U принадлежит множеству $\cup A'$. В то же время множества A и A' являются частями множества U. Следовательно, каждый элемент их объединения $A\cup A'$ принадлежит множеству U. По определению равенства множеств сказанное означает, что множества $A\cup A'$ и U равны. Второе равенство доказано.

3. Если x принадлежит множеству A, то он не принадлежит дополнению A' множества A: если $x \in A$, то $x \notin A'$. В то же время если x не принадлежит множеству A', то он принадлежит дополнению A'' множества A': если $x \notin A'$, то $x \in A''$. Таким образом, если $x \in A$, то $x \in A''$. Обратно: если $x \in A''$, то $x \notin A'$ и, следовательно, $x \in A$. Значит, множества A и A'' равны. Третье равенство доказано. +

Исследуем связь между объединением, пересечением и дополнением для множеств.

Пример. Если $A = \{1, 2, 3\}$, $B = \{2, 3, 4\}$, $U = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, то $(A \cup B)' = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\} - \{1, 2, 3, 4\} = \{0, 5, 6, 7, 8, 9\} = \{0, 4, 5, 6, 7, 8, 9\} \cap \{0, 1, 5, 6, 7, 8, 9\} = A' \cap B'$. $(A \cap B)' = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\} - \{2, 3\} = \{0, 4, 5, 6, 7, 8, 9\} = \{0, 4, 5, 6, 7, 8, 9\} \cup \{0, 2, 5, 6, 7, 8, 9\} = A' \cup B'$.

Равенства крайних множеств имеют общий характер: *дополнение объединения двух множеств равно пересечению их дополнений, а дополнение пересечения двух множеств равно объединению их дополнений.*

Рассмотрим множество U и его произвольные части A, B. Связь между объединением, пересечением и дополнением выражает

Теорема де Моргана. $(A \cup B)' = A' \cap B'$, $(A \cap B)' = A' \cup B'$.

— 1. Докажем первое из этих равенств. Рассмотрим произвольный элемент x множества U.

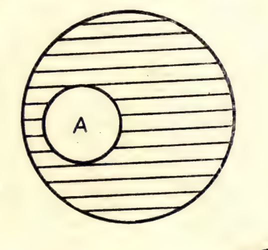
Если $x \in (A \cup B)'$, то по определению дополнения $x \notin A \cup B$. По определению объединения отсюда следует, что $x \notin A$ и $x \notin B$, т. е. $x \in A'$ и $x \in B'$. По определению пересечения это значит, что $x \in A' \cap B'$. Таким образом, если $x \in (A \cup B)'$, то $x \in A' \cap B'$.

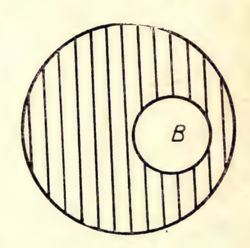
Обратно: если $x \in A' \cap B'$, то $x \in A'$ и $x \in B'$, т. е. $x \notin A$ и $x \notin B$. Отсюда следует, что $x \notin A \cup B$, т. е. $x \in (A \cup B)'$. Таким образом, если $x \in A' \cap B'$, то $x \in (A \cup B)'$.

По определению равенства множеств из сказанного следует, что множества $(A \cup B)'$ и $A' \cap B'$ равны. Первое равенство доказано.

2. Второе равенство следует из первого и из теоремы о дополнениях: $(A \cap B)' = (A'' \cap B'')' = ((A' \cup B')')' = (A' \cup B')'' = A' \cup B'$.

Теорема де Моргана доказана. +





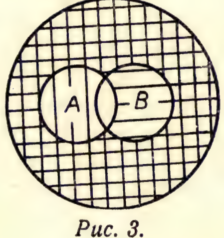


Рис. 3.

Доказательство поясняет рис. 3. Заштрихованная область на рисунке изображает объединение $A' \cup B'$, а дважды заштрихованная — пересечение $A' \cap B'$ дополнений A' и B'.

Теорема де Моргана позволяет заменять операцию объединения множеств операцией пересечения и наоборот. Иногда это бывает очень удобно.

§ 4*. КОЛЬЦА МНОЖЕСТВ

В этом параграфе определяются сложение и умножение множеств. Как сложение и умножение чисел, сложение и умножение множеств ассоциативны и коммутативны, причем умножение дистрибутивно относительно сложения. Сложение множеств совпадает со своей обратной операцией — вычитанием множеств.

Умножение множеств определяется так, что квадрат каждого множества равен самому этому множеству. Эти свойства отличают сложение и умножение множеств от сложения и умножения чисел.

1.4. Сложение множеств

Естественно предложить взять в качестве суммы $A+B$ множеств A, B их объединение $A \cup B$. Однако при таком определении суммы множеств не для всех множеств A и B будет определена их разность $B-A$. Действительно, если A не является частью B, то $X \cup A$ не является частью B и, следовательно, $X \cup A \neq B$ для каждого множества X.

Предлагается следующее

Определение суммы. *Множество*

$$A + B = (A \cup B) - (A \cap B)$$

называется суммой множеств A, B.

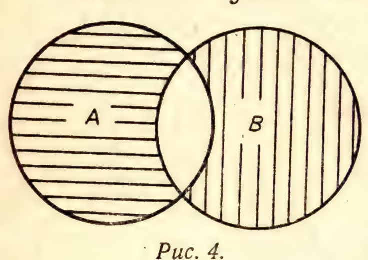


Рис. 4.

Определение суммы множеств поясняет рис. 4. На этом рисунке левый круг с горизонтальной штриховкой изображает множество A, правый круг с вертикальной штриховкой — множество B, а вся заштрихованная фигура — сумму $A+B$.

Можно сказать, что сумма множеств равна их объединению без пересечения.

Сумма множеств A, B называется также их *симметрической разностью* и обозначается символом $A \Delta B$.

Рассмотрим несколько примеров сумм множеств.

Пример 1. Если $A=\{1, 2, 3\}$, $B=\{2, 3, 4\}$, то $A+B= = (\{1, 2, 3\} \cup \{2, 3, 4\}) - (\{1, 2, 3\} \cap \{2, 3, 4\}) = \{1, 2, 3, 4\} - \{2, 3\} = = \{1, 4\}$.

Пример 2. Если $A=B$, то $A+A=(A\cup A)-(A\cap A)=A-A=\varnothing$.

Пример 3. Если $A\cap B=0$, то $A+B=(A\cup B)-(A\cap B)=A\cup B$.

Прежде, чем сформулировать и доказать теорему об основных свойствах сложения, докажем следующее предложение, описывающее свойства операций объединения и пересечения.

Лемма об объединении и пересечении. *Объединение и пересечение множеств коммутативны, ассоциативны и дистрибутивны друг относительно друга.*

— Высказанное предложение означает, что для каждых множеств A, B и C верны равенства

$$A\cup B=B\cup A,\quad A\cap B=B\cap A,$$
$$(A\cup B)\cup C=A\cup(B\cup C),\quad (A\cap B)\cap C=A\cap(B\cap C),$$
$$A\cap(B\cup C)=(A\cap B)\cup(A\cap C),\quad A\cup(B\cap C)=(A\cup B)\cap(A\cup C).$$

Вследствие теоремы де Моргана в каждой строке вторые равенства вытекают из первых. Например,

$$A\cap B=A''\cap B''=(A'\cup B')'=(B'\cup A')'=B''\cap A''=B\cap A.$$

Таким образом, для доказательства леммы достаточно доказать первые равенства в каждой строке.

1. Равенство $A\cup B=B\cup A$ вытекает непосредственно из определения объединения множеств.

2. Равенство $(A\cup B)\cup C=A\cup(B\cup C)$ верно, так как каждое из рассматриваемых множеств образовано элементами, принадлежащими множеству A или B, или C.

3. Если $x\in A\cap(B\cup C)$, то $x\in A$ и $x\in B\cup C$. Если $x\in B\cup C$, то $x\in B$ или $x\in C$. Следовательно, если $x\in A$ и $x\in B\cup C$, то а) $x\in A$ и $x\in B$ или б) $x\in A$ и $x\in C$, т. е. $x\in A\cap B$ или $x\in A\cap C$. Значит, если $x\in A\cap(B\cup C)$, то $x\in(A\cap B)\cup(A\cap C)$. Проведя это рассуждение в обратном порядке, убеждаемся в том, что, если $x\in(A\cap B)\cup(A\cap C)$, то $x\in A\cap(B\cup C)$. Таким образом, равенство $A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$ верно.

Лемма доказана. +

Основные свойства сложения множеств выражает

Теорема о сложении. *Сложение множеств коммутативно и ассоциативно.*

— 1. Коммутативность сложения для множеств означает, что для каждых множеств A и B верно равенство

$$A+B=B+A.$$

Это равенство вытекает из определения суммы множеств и коммутативности объединения и пересечения множеств:

$$A+B=(A\cup B)-(A\cap B)=(B\cup A)-(B\cap A)=B+A.$$

2. Ассоциативность сложения означает, что для каждых множеств A, B и C верно равенство

$$(A+B)+C=A+(B+C).$$

Это равенство верно, так как каждое из рассматриваемых множеств, как нетрудно проверить, образовано элементами, принадлежащими ровно одному из множеств A, B, C или всем трем, т. е.

$$(A+B)+C=(A\cap B'\cap C')\cup(A'\cap B\cap C')\cup(A'\cap B'\cap C)\cup$$
$$\cup(A\cap B\cap C)=A+(B+C)$$

(имеются в виду дополнения до множества $U=A\cup B\cup C$).

Теорема доказана. +

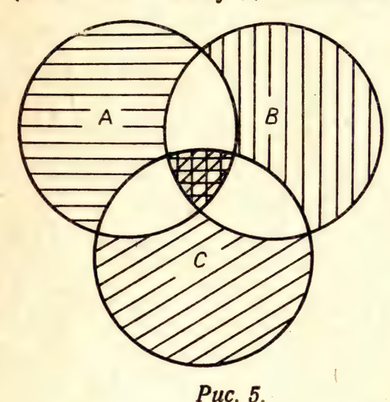


Рис. 5.

Доказательство поясняет рис. 5. На этом рисунке множества $A\cap B'\cap C'$, $A'\cap B\cap C'$ и $A'\cap B'\cap C$ изображаются соответственно горизонтально, вертикально и косо заштрихованными областями. Множество $A\cap B\cap C$ — трижды заштрихованная область.

Из определений следует, что пустое множество O является *нулем* для сложения множеств:

$$A+O=A$$

для каждого множества A. Поэтому можно говорить о множестве $-A$, *противоположном* A:

$$A+(-A)=O.$$

Лемма о противоположном множестве. *Множество $-A$, противоположное множеству A, равно множеству A:*

$$A+A=O,\quad -A=A.$$

— Действительно, как уже отмечалось в примере 2, для каждого множества A верны равенства

$$A+A=(A\cup A)-(A\cap A)=A-A=O.$$

Лемма доказана. +

Следствие. *Разность каждых множеств A и B равна их сумме*:

$$A-B=A+B.$$

— Действительно, используя ассоциативность сложения множеств и лемму о противоположном множестве, получаем

$$(A+B)+B=A+(B+B)=A+O=A.$$

Разность множеств A и B по определению есть множество, сумма которого с B равна A. Из полученных равенств следует, что эта разность равна $A+B$. Высказанное предложение доказано. +

Таким образом, вычитание множеств совпадает со сложением множеств.

2.4. Умножение множеств

В качестве умножения множеств выбирается их пересечение. Поэтому дается следующее

Определение произведения. *Множество* $A \cdot B = A \cap B$ *называется произведением множеств* A, B.

Основное отличие умножения множеств от умножения чисел выражает следующее предложение, которое вытекает непосредственно из определения умножения множеств.

Правило квадрата. *Квадрат каждого множества* A *равен множеству* A:

$$A^2 = A.$$

Основные свойства умножения множеств выражает

Теорема об умножении. *Умножение множеств коммутативно, ассоциативно и дистрибутивно относительно сложения.*

— Коммутативность и ассоциативность умножения следуют непосредственно из леммы об объединении и пересечении. Дистрибутивность умножения множеств относительно сложения означает, что для каждых множеств A, B и C верно равенство

$$A(B+C) = AB+AC.$$

Докажем это равенство. По определению

$$A(B+C) = A \cap [(B \cup C) - (B \cap C)].$$

Как нетрудно проверить,

$$A \cap [(B \cup C) - (B \cap C)] = A \cap (B \cup C) - A \cap (B \cap C).$$

Вследствие дистрибутивности пересечения относительно объединения

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Вследствие правила квадрата, коммутативности и ассоциативности пересечения

$$A \cap (B \cap C) = (A \cap A) \cap (B \cap C) = (A \cap B) \cap (A \cap C).$$

Из определений вытекает, что

$$[(A \cap B) \cup (A \cap C)] - [(A \cap B) \cap (A \cap C)] = AB + AC.$$

Непосредственно из полученных равенств следует равенство

$$A(B+C) = AB+AC.$$

Теорема доказана. +

Доказательство поясняет рис. 6. На этом рисунке множества A, $B+C$ и $A(B+C)$ изображаются соответственно областями с горизонтальной, вертикальной и двойной штриховкой.

Теоремы о сложении и умножении множеств выражают свойства этих операций, аналогичные свойствам сложения и умноже-

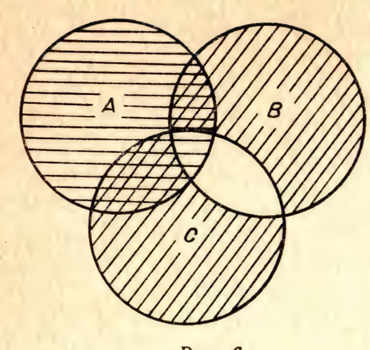


Рис. 6.

ния чисел. Лемма о противоположном множестве и правило квадрата показывают существенную разницу между свойствами операций для множеств и для чисел. Для каждого множества A верны равенства

$$A+A=O, \quad A\cdot A=A.$$

В то же время для каждого числа a

$$a+a\neq 0, \quad a\cdot a\neq a,$$

если, соответственно, $a\neq 0$ и $a\neq 0$, 1.

3.4. Формулы перехода

Сумма и произведение множеств выражаются через их объединение, пересечение и дополнение:

$$(1)\quad A+B=(A\cup B)-(A\cap B), \quad A\cdot B=A\cap B.$$

В свою очередь, объединение, пересечение и дополнение множеств могут быть выражены через их сумму и произведение: непосредственно из определений следует, что для каждых множеств A и B верны равенства

$$(2)\quad A\cup B=A+B+A\cdot B, \quad A\cap B=A\cdot B.$$

Если множество A является частью множества B, то для дополнения $B-A$ множества A верно равенство

$$(3)\quad B-A=A+B.$$

Условимся равенства (1), (2) и (3) называть формулами перехода от объединения, пересечения и дополнения к сложению и умножению множеств и обратно.

4.4. Кольца множеств

Часто приходится рассматривать множества, элементы которых сами являются множествами. Условимся такие *множества множеств* называть *классами* множеств. Для каждого множества U существует класс $\mathscr{P}$, образованный всеми частями множества U. Каждая часть класса $\mathscr{P}$ также является классом множеств.

Пример 1. Если U является пустым множеством, то класс $\mathscr{P}$ образован одним единственным множеством — пустым множеством O:

$$\mathscr{P}=\{O\}.$$

Пример 2. Если U является элементарным множеством, то класс $\mathscr{P}$ образован двумя множествами — пустым множеством и множеством U:

$$\mathscr{P}=\{O, U\}.$$

Пример 3. Если $U = D = \{0, 1\}$, то класс $\mathscr{P}$ образован четырьмя множествами — пустым множеством O, элементарными множествами $\{0\}$, $\{1\}$ и множеством D:

$$\mathscr{P} = \{O, \{0\}, \{1\}, D\}.$$

Условимся говорить, что для класса $\mathscr{A}$ множеств определено сложение, если сумма каждых множеств класса $\mathscr{A}$ принадлежит этому классу. Точно так же будем говорить, что для класса $\mathscr{A}$ множеств определено умножение, если произведение каждых двух множеств класса $\mathscr{A}$ принадлежит этому классу.

Пример 1. Если $\mathscr{A} = \mathscr{P}$, то для $\mathscr{A}$ определены сложение и умножение.

Пример 2. Рассмотрим класс $\mathscr{P}$ всех частей множества $U = \{00, 01, 10, 11\}$, элементами которого являются пары, составленные из чисел 0 и 1: 00 (*ноль, ноль*), 01 (*ноль, один*), 10 (*один, ноль*), 11 (*один, один*). Элементами класса $\mathscr{P}$ являются, в частности, множества $H = \{00, 01\}$ и $У = \{10, 11\}$. Для класса $\mathscr{A} = \{O, H, У, U\}$, образованного множествами O, H, $У$, U, определены сложение и умножение:

Таблица сложения

+	O	H	У	U
O	O	H	У	U
H	H	O	U	У
У	У	U	O	H
U	U	У	H	O

Таблица умножения

·	O	H	У	U
O	O	O	O	O
H	O	H	O	H
У	O	O	У	У
U	O	H	У	U

На пересечении строки X и столбца Y в таблице сложения помещена их сумма $X + Y$, а в таблице умножения — произведение $X \cdot Y$ множеств класса $\mathscr{A}$.

Контрпример. Для класса $\mathscr{A} = \{\{0\}, \{1\}\}$ частей множества $D = \{0, 1\}$ не определены ни сложение, ни умножение: множества $D = \{0\} + \{1\}$ и $O = \{0\} \cdot \{1\}$ не принадлежат классу $\mathscr{A}$.

Классы множеств, для которых определены сложение и умножение, играют важную роль в теории множеств. Поэтому для них вводится специальный термин.

Определение кольца множеств. *Если для непустого класса $\mathscr{A}$ множеств определены сложение и умножение, то говорят, что класс $\mathscr{A}$ образует кольцо.*

Примеры 1 и 2, предпосланные определению, являются примерами классов множеств, образующих кольца. Контрпример является примером класса, не образующего кольца.

5.4. Свойства колец множеств

Общие свойства колец множеств вытекают из ассоциативности сложения и умножения множеств, а также дистрибутивности умножения относительно сложения.

Отметим следующие свойства колец множеств.

Рассмотрим произвольный непустой класс $\mathscr{A}$ множеств, образующий кольцо.

Свойство 1. *Пустое множество O принадлежит классу $\mathscr{A}$ и является нулем образованного классом $\mathscr{A}$ кольца множеств.*

Рассмотрим произвольное множество A класса $\mathscr{A}$. Так как $A+A=O$, то O принадлежит классу $\mathscr{A}$. Следующее из определений равенство

$$A+O=A$$

означает, что пустое множество O является нулем кольца множеств, образованного классом $\mathscr{A}$. +

Для каждого множества A верно также равенство

$$A \cdot O=O.$$

Предположим дополнительно, что класс $\mathscr{A}$ составлен из некоторых частей множества U. По определению, множество U является *единицей* образованного классом $\mathscr{A}$ кольца, если равенство

$$A \cdot U=A$$

верно для каждого множества A класса $\mathscr{A}$.

Свойство 2. *Если множество U принадлежит классу $\mathscr{A}$, то оно является единицей образованного классом $\mathscr{A}$ кольца.*

— Равенство

$$A \cdot U=A$$

верно для каждой части A множества U. В частности, оно верно для каждого множества A класса $\mathscr{A}$. Если U принадлежит $\mathscr{A}$, то это и означает, что U является единицей образованного классом $\mathscr{A}$ кольца. +

Условимся говорить, что для класса $\mathscr{A}$ множеств определены объединение и пересечение, если объединение и пересечение каждых двух множеств класса $\mathscr{A}$ принадлежат этому классу. Будем говорить, что для класса $\mathscr{A}$ множеств определено дополнение, если для каждой принадлежащей классу $\mathscr{A}$ части A множества B класса $\mathscr{A}$ дополнение $B-A$ множества A до множества B принадлежит классу $\mathscr{A}$.

Свойство 3. *Если класс $\mathscr{A}$ множеств образует кольцо, то для класса $\mathscr{A}$ определены объединение, пересечение и дополнение.*

— Если класс $\mathscr{A}$ множеств образует кольцо, то для него определены сложение и умножение. Формулы перехода (2) и (3) показывают, что в этом случае для класса $\mathscr{A}$ определены объединение, пересечение и дополнение. +

Из формул перехода (1) вытекает

Свойство 4. *Если для класса $\mathscr{A}$ множеств определены объединение, пересечение и дополнение, то он образует кольцо.*

Таким образом, для класса $\mathscr{A}$ множеств, образующего кольцо, определены все рассматривавшиеся операции: объединение, пересечение, дополнение, сложение и умножение.

6.4. Пример

Рассмотрим класс $\mathscr{D}$, образованный множествами O и U. Условимся для большей выразительности обозначать множество U также символом 1. Таким образом,

$$\mathscr{D}=\{O, 1\}.$$

Класс $\mathscr{D}$ образует кольцо. Сложение и умножение для $\mathscr{D}$ определяют

Таблица сложения

+	0	1
0	0	1
1	1	0

Таблица умножения

·	0	1
0	0	0
1	0	1

На пересечении строки X и столбца Y в таблице сложения помещена сумма $X+Y$, а в таблице умножения — произведение $X \cdot Y$ множеств X и Y класса $\mathscr{D}$. Множества O и 1 являются соответственно нулем и единицей кольца, образованного классом $\mathscr{D}$. Таблица сложения для множеств O и 1 отличается от таблицы сложения для чисел 0 и 1 равенством

$$1+1=O.$$

Правило «*один плюс один равно ноль*» выражает специфику сложения множеств. Таблица умножения для множеств O и 1 аналогична таблице умножения для чисел 0 и 1.

Объединение и пересечение для класса $\mathscr{D}$ определяют

Таблица объединения

∪	0	1
0	0	1
1	1	1

Таблица пересечения

∩	0	1
0	0	0
1	0	1

На пересечении строки X и столбца Y в таблице объединения помещено объединение, а в таблице пересечения — пересечение

множеств X и Y класса $\mathscr{D}$. Таблица объединения отличается от таблицы сложения равенством

$$1 \cup 1 = 1.$$

Таблица пересечения, естественно, совпадает с таблицей умножения.

Если для класса $\mathscr{D}$ в качестве отношения порядка взять отношение включения, т. е. считать, что O строго меньше 1:

$$O < 1,$$

то это отношение порядка будет аналогично обычному отношению порядка для чисел 0 и 1. При этом для каждых множеств X и Y класса $\mathscr{D}$ будут верны равенства

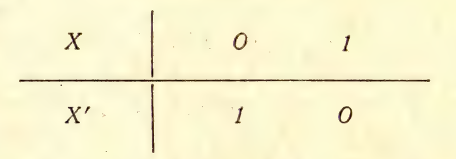
$$X \cup Y = \max\{X, Y\}, \quad X \cap Y = \min\{X, Y\}.$$

Таблицы объединения и пересечения совпадают в этом случае с таблицами максимума и минимума для множеств и аналогичны таблицам максимума и минимума для чисел 0 и 1. Равенство для объединения множества 1 с самим собой принимает привычный вид

$$\max\{1, 1\} = 1.$$

Дополнение для класса $\mathscr{D}$ определяет

Таблица дополнения



X	O	1
X'	1	O

В этой таблице под множеством X класса $\mathscr{D}$ помещено его дополнение X' до множества $U = 1$. Это дополнение получается с помощью равенства

$$X' = X + 1.$$

Замечание. Класс $\mathscr{D}$ с так определенными операциями объединения, пересечения и дополнения является простейшим примером *булевой алгебры.* Булевы алгебры играют важную роль в некоторых разделах математики. (Подробнее по поводу булевых алгебр см., например, книгу Дж. Т. Калбертсона «Математика и логика цифровых устройств». М., «Просвещение», 1965.)

§ 5*. ДЕКАРТОВЫ ПРОИЗВЕДЕНИЯ

Часто приходится рассматривать упорядоченные пары, образованные элементами двух множеств, и различные множества упорядоченных пар. Понятие *упорядоченной пары* представляется

интуитивно ясным: отец и сын, мать и дочь образуют пары, упорядоченные по старшинству. Используя понятие множества, можно дать формальное определение упорядоченной пары.

1.5. Определение упорядоченной пары

Рассмотрим произвольный элемент a множества A и произвольный элемент b множества B.

Определение упорядоченной пары. *Класс*

$$(a, b) = \{\{a\}, \{a, b\}\},$$

составленный из множеств $\{a\}$ и $\{a, b\}$, называется упорядоченной парой, образованной элементами a и b.

Элемент a называется *первым* элементом, а элемент b — *вторым* элементом упорядоченной пары (a, b). Вместо (a, b) часто пишут ab.

Пример 1. Если $A = B = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, буквы a и b обозначают произвольные цифры, то упорядоченной парой ab является двузначное число: 00, 01, ..., 09, 10, 11, ..., 19, ..., 90, 91, ..., 99.

Заметим, что $(a, b) \neq (b, a)$, если $a \neq b$. Порядок, в котором записываются элементы a и b, существенен. Этим упорядоченная пара (a, b) отличается от множества $\{a, b\}$, образованного элементами a и b.

Пример 2. Если каждое из множеств A и B равно множеству R вещественных чисел, буквы a и b обозначают произвольные вещественные числа, то упорядоченную пару (a, b) называют точкой вещественной координатной плоскости. Числа a и b называются соответственно первой и второй координатами точки (a, b) (рис. 7).

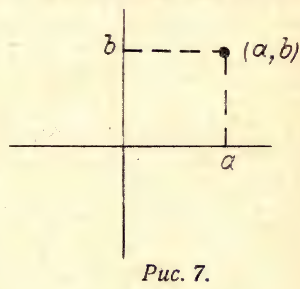


Рис. 7.

Пример 3. Если $A = B = \{0, 1\}$, то множество всех упорядоченных пар (a, b) образуется упорядоченными парами (0, 0), (0, 1), (1, 0), (1, 1).

2.5. Равенство упорядоченных пар

Рассмотрим упорядоченные пары (a, b) и (c, d), образованные произвольными элементами a, c множества A и элементами b, d множества B.

Теорема о равенстве упорядоченных пар. *Равенство $(a, b) = (c, d)$ упорядоченных пар (a, b) и (c, d) эквивалентно равенству $a = c$ первых элементов и равенству $b = d$ вторых элементов этих пар.*

— 1. Если $a = c$, $b = d$, то

$$(a, b) = \{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\} = (c, d).$$

2. Если $(a, b)=(c, d)$, то $\{a\}=\{c\}$ и $\{a, b\}=\{c, d\}$ или $\{a\}=\{c, d\}$ и $\{a, b\}=\{c\}$.

В первом случае из равенства $\{a\}=\{c\}$ вытекает равенство $a=c$, а из этого равенства и равенства $\{a, b\}=\{c, d\}$ следует равенство $b=d$. Таким образом, в первом случае $a=c$ и $b=d$.

Во втором случае из равенства $\{a\}=\{c, d\}$ вытекают равенства $a=c=d$, а из равенства $\{a, b\}=\{c\}$ следуют равенства $a=b=c$. Таким образом, и во втором случае $a=c$ и $b=d$. Следовательно, если $(a, b)=(c, d)$, то $a=c$ и $b=d$.

Теорема доказана. +

Если $A=B$, то при $b=c$ и $a=d$ непосредственно из теоремы о равенстве упорядоченных пар вытекает

Следствие. *Равенство* $(a, b)=(b, a)$ *упорядоченных пар* (a, b) и (b, a) *эквивалетно равенству* $a=b$.

Говорят, что упорядоченная пара (b, a) симметрична упорядоченной паре (a, b).

3.5. Определение декартова произведения

Для каждых множеств A и B существует множество $A\times B$, образованное всеми упорядоченными парами (a, b), первыми элементами которых являются элементы a множества A, а вторыми — элементы b множества B:

$$A\times B=\{(a, b)/a\in A, \quad b\in B\}.$$

Определение декартова произведения. *Декартовым произведением множества A на множество B называется множество $A\times B$, образованное всеми упорядоченными парами, первые элементы которых принадлежат множеству A, а вторые — множеству B.*

Если множества A и B различны, то декартово произведение $A\times B$ не равно декартову произведению $B\times A$.

§ 6*. ОТНОШЕНИЯ И ОТОБРАЖЕНИЯ

Содержание понятий *отношения* и *отображения* поясняется употреблением этих терминов в обычной речи: отношение родства, образ предмета. Используя понятие декартова произведения, можно дать математические определения понятий отношения и отображения.

1.6. Определение отношения

Прежде чем сформулировать определение отношения, рассмотрим несколько простых примеров.

Рассмотрим множества $A=B=\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ и их декартово произведение

$$A\times B=\left\{\begin{matrix} 00 & 01 & 02 & 03 & 04 & 05 & 06 & 07 & 08 & 09 \\ 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 \\ 20 & 21 & 22 & 23 & 24 & 25 & 26 & 27 & 28 & 29 \\ 30 & 31 & 32 & 33 & 34 & 35 & 36 & 37 & 38 & 39 \\ 40 & 41 & 42 & 43 & 44 & 45 & 46 & 47 & 48 & 49 \\ 50 & 51 & 52 & 53 & 54 & 55 & 56 & 57 & 58 & 59 \\ 60 & 61 & 62 & 63 & 64 & 65 & 66 & 67 & 68 & 69 \\ 70 & 71 & 72 & 73 & 74 & 75 & 76 & 77 & 78 & 79 \\ 80 & 81 & 82 & 83 & 84 & 85 & 86 & 87 & 88 & 89 \\ 90 & 91 & 92 & 93 & 94 & 95 & 96 & 97 & 98 & 99 \end{matrix}\right\}$$

Пример 1. Отношение равенства для цифр определяется частью **e** множества $A\times B$:

$$\mathbf{e}=\left\{\begin{matrix} 00 & & & & & & & & & \\ & 11 & & & & & & & & \\ & & 22 & & & & & & & \\ & & & 33 & & & & & & \\ & & & & 44 & & & & & \\ & & & & & 55 & & & & \\ & & & & & & 66 & & & \\ & & & & & & & 77 & & \\ & & & & & & & & 88 & \\ & & & & & & & & & 99 \end{matrix}\right\}$$

Равенство $a=b$ означает, что упорядоченная пара ab принадлежит множеству **e**.

Пример 2. Отношение порядка для цифр определяется частью **o** множества $A\times B$:

$$\mathbf{o}=\left\{\begin{matrix} 00 & 01 & 02 & 03 & 04 & 05 & 06 & 07 & 08 & 09 \\ & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 \\ & & 22 & 23 & 24 & 25 & 26 & 27 & 28 & 29 \\ & & & 33 & 34 & 35 & 36 & 37 & 38 & 39 \\ & & & & 44 & 45 & 46 & 47 & 48 & 49 \\ & & & & & 55 & 56 & 57 & 58 & 59 \\ & & & & & & 66 & 67 & 68 & 69 \\ & & & & & & & 77 & 78 & 79 \\ & & & & & & & & 88 & 89 \\ & & & & & & & & & 99 \end{matrix}\right\}$$

Неравенство $a\leqslant b$ означает, что упорядоченная пара ab принадлежит множеству **o**.

Пример 3. Отношение строгого порядка для цифр определяется частью **s** множества $A \times B$:

$$\mathbf{s} = \left\{ \begin{array}{ccccccccc} 01 & 02 & 03 & 04 & 05 & 06 & 07 & 08 & 09 \\ & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 \\ & & 23 & 24 & 25 & 26 & 27 & 28 & 29 \\ & & & 34 & 35 & 36 & 37 & 38 & 39 \\ & & & & 45 & 46 & 47 & 48 & 49 \\ & & & & & 56 & 57 & 58 & 59 \\ & & & & & & 67 & 68 & 69 \\ & & & & & & & 78 & 79 \\ & & & & & & & & 89 \end{array} \right\}$$

Строгое неравенство $a < b$ означает, что упорядоченная пара ab принадлежит множеству **s**.

Пример 4. Рассмотрим произведение $A \times B$ множества $A = \{0, 2, 4, 6, 8\}$ на множество $B = \{1, 3, 5, 7, 9\}$. Отношение «*a непосредственно предшествует b*» определяется частью

$$\mathbf{r} = \{01, 23, 45, 67, 89\}$$

множества $A \times B$. Это отношение означает, что упорядоченная пара ab принадлежит множеству **r**.

Примеры 1—4 поясняют

Определение отношения. *Каждая часть* **r** *множества* $A \times B$ *называется отношением между множеством A и множеством B.*

Определение отношения как множества упорядоченных пар позволяет просто описывать и наглядно представлять многие свойства отношений.

Вместо «упорядоченная пара (a, b), образованная элементами a и b, принадлежит отношению **r**» условимся также говорить «элемент *a связан отношением* **r** с элементом b». Именно в этом смысле в примерах 1, 2 и 3 обычно употребляются выражения «a равно b», «a меньше b» и «a строго меньше b».

2.6. Образы и прообразы

Рассмотрим произвольное отношение **r** между множеством A и множеством B.

Для каждого элемента $a \in A$ множество

$$\mathbf{r}(a) = \{b / (a, b) \in \mathbf{r}\}$$

всех элементов $b \in B$, для которых упорядоченная пара (a, b) принадлежит отношению **r**, называется *образом элемента a* при отношении **r**.

Пример 1. При отношении равенства $\mathbf{e}(a) = \{a\}$ для каждого элемента a множества $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

Пример 2. При отношении порядка $\mathbf{o}(9) = \{9\}$, $\mathbf{o}(8) = \{8, 9\}$, $\mathbf{o}(7) = \{7, 8, 9\}, \ldots$

Пример 3. При отношении строгого порядка $\mathbf{s}(9) = O$, $\mathbf{s}(8) = \{9\}$, $\mathbf{s}(7) = \{8, 9\}$, ...

Пример 4. При отношении «непосредственно предшествует» $\mathbf{r}(0) = \{1\}$, $\mathbf{r}(2) = \{3\}$, ...

Пример 3 показывает, что образ элемента может быть пустым.

Для каждого множества $X \subseteq A$ множество $\mathbf{r}(X)$ всех элементов b из B, принадлежащих образу $\mathbf{r}(a)$ хотя бы одного элемента a из X, называется *образом множества* X при отношении $\mathbf{r}$. Образ множества A называется *областью значений* отношения $\mathbf{r}$.

Пример 1. Областью значений отношения равенства для цифр является множество всех цифр.

Пример 2. Областью значений отношения порядка для цифр является множество всех цифр.

Пример 3. Областью значений отношения строгого порядка для цифр является множество $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Цифра 0 не принадлежит этой области, так как не существует цифры, строго меньшей цифры 0.

Пример 4. Областью значений отношения *непосредственно предшествует* между четными и нечетными цифрами является множество нечетных цифр.

Пример 3 показывает, что область значений $\mathbf{r}(A)$ отношения $\mathbf{r}$ между множеством A и множеством B может быть собственной частью множества B.

Для каждого элемента $b \in B$ множество

$$\mathbf{r}^{-1}(b) = \{a / (a, b) \in \mathbf{r}\}$$

всех элементов $a \in A$, для которых упорядоченная пара (a, b) принадлежит отношению $\mathbf{r}$, называется *прообразом элемента* b при отношении $\mathbf{r}$.

Пример 1. При отношении равенства $\mathbf{e}^{-1}(b) = \{b\}$ для каждого элемента b множества $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

Пример 2. При отношении порядка $\mathbf{o}^{-1}(0) = \{0\}$, $\mathbf{o}^{-1}(1) = \{0, 1\}$, $\mathbf{o}^{-1}(2) = \{0, 1, 2\}$, ...

Пример 3. При отношении строгого порядка $\mathbf{s}^{-1}(0) = O$, $\mathbf{s}^{-1}(1) = \{0\}$, $\mathbf{s}^{-1}(2) = \{0, 1\}$, ...

Пример 4. При отношении «непосредственно предшествует» $\mathbf{r}^{-1}(1) = \{0\}$, $\mathbf{r}^{-1}(3) = \{2\}$, $\mathbf{r}^{-1}(5) = \{4\}$, ...

Пример 3 показывает, что прообраз элемента может быть пустым множеством.

Для каждого множества $Y \subseteq B$ множество $\mathbf{r}^{-1}(Y)$ всех элементов a из A, принадлежащих прообразу $\mathbf{r}^{-1}(b)$ хотя бы одного элемента b из Y, называется *прообразом множества* Y при отношении $\mathbf{r}$. Прообраз $\mathbf{r}^{-1}(B)$ множества B называется *областью определения* отношения $\mathbf{r}$.

Пример 1. Областью определения отношения равенства для цифр является множество всех цифр.

Пример 2. Областью определения отношения порядка для цифр является множество всех цифр.

Пример 3. Областью определения отношения строгого порядка для цифр является множество $\{0, 1, 2, 3, 4, 5, 6, 7, 8\}$. Цифра 9 не принадлежит этой области, так как не существует цифры, которой цифра 9 строго меньше.

Пример 4. Областью определения отношения *непосредственно предшествует* между четными и нечетными цифрами является множество четных цифр.

Пример 3 показывает, что область определения $\mathbf{r}^{-1}(B)$ отношения $\mathbf{r}$ между множеством A и множеством B может быть собственной частью множества A.

Понятия образов и прообразов соответственно поясняют рис. 8 и 9.

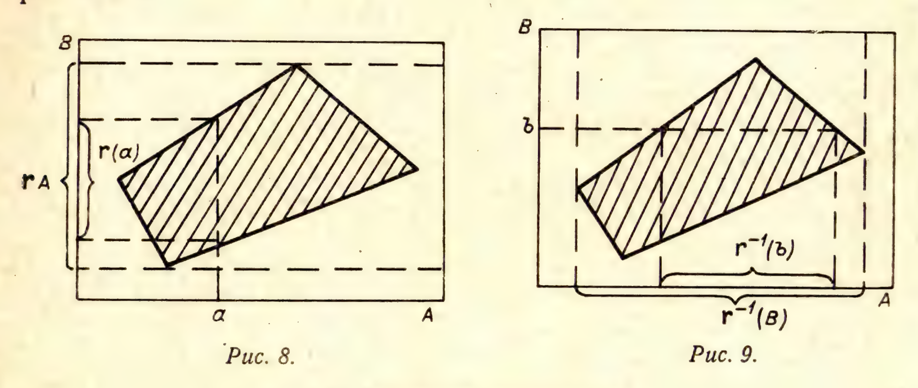


Рис. 8. *Рис. 9.*

3.6. Обратное отношение

Рассмотрим произвольное отношение $\mathbf{r}$ между множеством A и множеством B. Отношение

$$\mathbf{r}^{-1} = \{(b, a)/(a, b) \in \mathbf{r}\},$$

образованное всеми упорядоченными парами $(b, a) \in B \times A$, для которых симметричная упорядоченная пара (a, b) принадлежит отношению $\mathbf{r}$, называется *обратным* отношению $\mathbf{r}$.

Пример 1. Обратным отношению равенства является оно само: $\mathbf{e}^{-1} = \mathbf{e}$.

Пример 2. Обратным отношению порядка «$\leqslant$» является отношение порядка «$\geqslant$»:

$$\mathbf{o}^{-1} = \left\{\begin{matrix} 00 \\ 10 & 11 \\ 20 & 21 & 22 \\ 30 & 31 & 32 & 33 \\ 40 & 41 & 42 & 43 & 44 \\ 50 & 51 & 52 & 53 & 54 & 55 \\ 60 & 61 & 62 & 63 & 64 & 65 & 66 \\ 70 & 71 & 72 & 73 & 74 & 75 & 76 & 77 \\ 80 & 81 & 82 & 83 & 84 & 85 & 86 & 87 & 88 \\ 90 & 91 & 92 & 93 & 94 & 95 & 96 & 97 & 98 & 99 \end{matrix}\right\}$$

Неравенство $b \geqslant a$ означает, что упорядоченная пара $(b, a) = ba$ принадлежит множеству $\mathbf{o}^{-1}$, т. е. что $(a, b) = ab \in \mathbf{o}$ и $a \leqslant b$.

Пример 3. Обратным отношению строгого порядка «$<$» является отношение строгого порядка «$>$»:

$$\mathbf{s}^{-1} = \left\{ \begin{array}{lllllllll} 10 \\ 20 & 21 \\ 30 & 31 & 32 \\ 40 & 41 & 42 & 43 \\ 50 & 51 & 52 & 53 & 54 \\ 60 & 61 & 62 & 63 & 64 & 65 \\ 70 & 71 & 72 & 73 & 74 & 75 & 76 \\ 80 & 81 & 82 & 83 & 84 & 85 & 86 & 87 \\ 90 & 91 & 92 & 93 & 94 & 95 & 96 & 97 & 98 \end{array} \right\}$$

Строгое неравенство $b > a$ означает, что упорядоченная пара $ba = (b, a)$ принадлежит множеству $\mathbf{s}^{-1}$, т. е. что $ab = (a, b) \in \mathbf{s}$ и $a < b$.

Пример 4. Обратным отношению *непосредственно предшествует* является отношение *непосредственно следует*:

$$\mathbf{r}^{-1} = \{10,\ 32,\ 54,\ 76,\ 98\}.$$

Отношение «b непосредственно следует за a» означает, что упорядоченная пара ba принадлежит множеству $\mathbf{r}^{-1}$, т. е. что $ab \in \mathbf{r}$ и a непосредственно предшествует b.

Из определений следует, что образ $\mathbf{r}^{-1}(b)$ элемента b и образ $\mathbf{r}^{-1}(Y)$ множества Y при обратном отношении $\mathbf{r}^{-1}$ являются соответственно прообразом элемента b и прообразом множества Y при отношении $\mathbf{r}$. Область значений $\mathbf{r}^{-1}(B)$ обратного отношения $\mathbf{r}^{-1}$ является областью определения отношения $\mathbf{r}$, а область значений $\mathbf{r}(A)$ отношения $\mathbf{r}$ — областью определения обратного отношения $\mathbf{r}^{-1}$.

Понятие обратного отношения поясняет рис. 10.

Если $A = B$, то отношения $\mathbf{r}$ и $\mathbf{r}^{-1}$ изображаются симметричными относительно диагонали $\mathbf{e} = \{(a, a) / a \in A\}$ областями в квадрате $A \times A$.

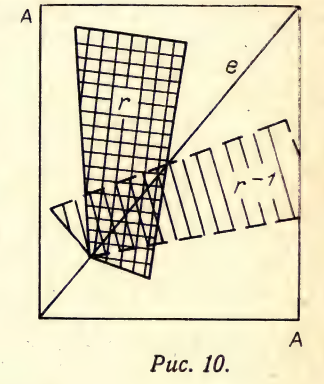


Рис. 10.

4.6. Композиция отношений

Рассмотрим непустые множества A, B и C. Упорядоченную пару $(a, c) \in A \times C$ назовем композицией упорядоченных пар $(a, b) \in A \times B$ и $(b, c) \in B \times C$. Для обозначения композиции условимся использовать символ $\circ$:

$$(a, c) = (a, b) \circ (b, c).$$

Будем говорить также, что упорядоченная пара (a, b) компонуется с упорядоченной парой (b, c). Если $b_1 \neq b_2$, то упорядоченная пара (a, b_1) не компонуется с упорядоченной парой (b_2, c).

Примеры. $(0, 2) = (0, 1) \circ (1, 2)$; $(1, 3) = (1, 2) \circ (2, 3)$; $(5, 9) = (5, 7) \cdot (7, 9)$; $(8, 1) = (8, 3) \cdot (3, 1)$; $(5, 5) = (5, 5) \cdot (5, 5)$; $(0, 1)$ не компонуется с $(2, 3)$; $(2, 3)$ не компонуется с $(4, 5)$.

Рассмотрим отношение **r** между множеством A и множеством B, отношение **s** между множеством B и множеством C. Отношение **s∘r** между множеством A и множеством C, составленное из всех возможных композиций упорядоченных пар отношения **r** с упорядоченными парами отношения **s**, будем называть композицией отношения **r** с отношением **s**.

Пример 1. Композиция отношения равенства **e** с самим собой снова дает отношение равенства **e∘e=e**, так как $(a, a) \circ (a, a) = (a, a)$ для каждой цифры a: $(0, 0) \circ (0, 0) = (0, 0)$, $(1, 1) \circ (1, 1) = (1, 1)$, ...

Пример 2. Композиция отношения порядка **o** с самим собой снова дает отношение порядка **o∘o=o**, так как композиция (a, c) каждой упорядоченной пары (a, b) отношения **o** с каждой парой (b, c) этого отношения принадлежит отношению **o**, а кроме того, каждая пара (a, c) отношения **o** является композицией упорядоченной пары (a, a) отношения **o** с упорядоченной парой (a, c) этого отношения. Эти утверждения проверяются непосредственно по таблице, определяющей отношение **o**:

$$(0, 0) \circ (0, 0) = (0, 0),\ (0, 0) \circ (0, 1) = (0, 1),$$
$$(0, 0) \circ (0, 2) = (0, 2),\ (1, 1) \circ (1, 1) = (1, 1), \ldots .$$

Пример 3. Композиция отношения строгого порядка **s** с самим собой строго содержится в **s**:

$$\mathbf{s} \circ \mathbf{s} = \left\{ \begin{matrix} 02 & 03 & 04 & 05 & 06 & 07 & 08 & 09 \\ & 13 & 14 & 15 & 16 & 17 & 18 & 19 \\ & & 24 & 25 & 26 & 27 & 28 & 29 \\ & & & 35 & 36 & 37 & 38 & 39 \\ & & & & 46 & 47 & 48 & 49 \\ & & & & & 57 & 58 & 59 \\ & & & & & & 68 & 69 \\ & & & & & & & 79 \end{matrix} \right\} \subset \mathbf{s}$$

Равенство для композиции **s∘s** проверяется по таблице, определяющей отношение **s**:

$$(0, 1) \cdot (1, 2) = (0, 2),\ (0, 1) \cdot (1, 3), \ldots, (1, 2) \cdot (2,3) = (1, 3), \ldots$$

Пример 4. Композиция отношения **r** непосредственного предшествования между четными и нечетными цифрами с самим собой является пустым отношением, так как упорядоченная пара, имеющая нечетную вторую цифру, не компонуется с упорядоченной парой, имеющей четную первую цифру.

Вместо «композиция отношений» говорят также «сложное отношение», сложные отношения между A и C выражаются через отношения между A и B и отношения между B и C.

5.6. Рефлексивность, симметричность и транзитивность

Эти три свойства отношений часто встречаются.

Произвольное отношение между множеством A и множеством $B=A$ условимся называть отношением для множества A. Множество

$$\mathbf{e}=\{(a, a)/a\in A\},$$

образованное всеми упорядоченными парами (a, a) элементов a множества A, условимся называть диагональю квадрата $A\times A$ множества A.

Отношение $\mathbf{r}$ для множества A называется *рефлексивным*, если каждый элемент a множества A связан отношением $\mathbf{r}$ с самим собой. (Т. е. если отношение $\mathbf{r}$ содержит диагональ $\mathbf{e}$ квадрата $A\times A$.)

Содержательный смысл рефлексивности поясняет отношение *знакомства*: каждый знаком с самим собой. Определение рефлексивности поясняет также рис. 11.

Пример 1. Отношение равенства $\mathbf{e}$ рефлексивно: $a=a$.

Пример 2. Отношение порядка $\mathbf{o}$ рефлексивно: $a\leqslant a$.

Пример 3. Отношение строгого порядка $\mathbf{s}$ не рефлексивно: $a\nless a$.

Отношение $\mathbf{r}$ для множества A называется *симметричным*, когда $\mathbf{r}$ обладает свойством: если элемент a связан отношением $\mathbf{r}$ с элементом b, то и элемент b, в свою очередь, связан отношением $\mathbf{r}$ с элементом a. (Т. е. если отношение $\mathbf{r}$ равно обратному отношению $\mathbf{r}^{-1}$.)

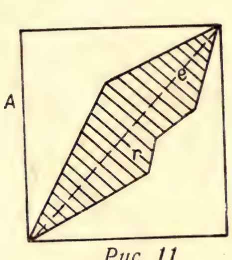


Рис. 11.

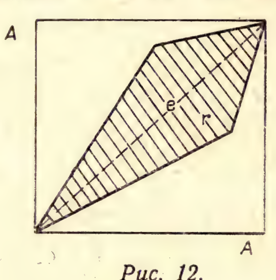


Рис. 12.

Содержание симметричности поясняет отношение *родства*: если a родственник b, то и b родственник a. Определение симметричности поясняет также рис. 12.

Симметричное отношение $\mathbf{r}$ изображается на этом рисунке областью, симметричной относительно диагонали квадрата.

Пример 1. Отношение равенства $\mathbf{e}$ симметрично: $\mathbf{e}=\mathbf{e}^{-1}$.

Пример 2. Отношение порядка $\mathbf{o}$ не симметрично: $\mathbf{o}\neq\mathbf{o}^{-1}$.

Пример 3. Отношение строгого порядка s не симметрично.

Отношение r для множества A называется *транзитивным*, когда r обладает свойством: если элемент a связан отношением r с элементом b и элемент b связан отношением r с элементом c, то элемент a связан отношением r с элементом c. (Т. е. если композиция $\mathbf{r} \circ \mathbf{r}$ отношения r с самим собой содержится в r.)

Содержательный смысл условия транзитивности поясняет следующий пример: если город a связан железной дорогой с городом b и город b связан железной дорогой с городом c, то город a связан железной дорогой с городом c.

Транзитный пассажир железной дороги попадает из города a в город c через город b.

Пример 1. Отношение равенства e транзитивно: $\mathbf{e} \circ \mathbf{e} = \mathbf{e}$; если $a = b$ и $b = c$, то $a = c$.

Пример 2. Отношение порядка o транзитивно: $\mathbf{o} \circ \mathbf{o} = \mathbf{o}$; если $a \leqslant b$ и $b \leqslant c$, то $a \leqslant c$.

Пример 3. Отношение строгого порядка транзитивно: $\mathbf{s} \circ \mathbf{s} \subseteq \mathbf{s}$; если $a < b$ и $b < c$, то $a < c$.

Пример нетранзитивного отношения. Для множества $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ цифр определим отношение *непосредственного соседства*:

$$\mathbf{r} = \left\{ \begin{array}{l} 01 \\ 10 \quad 12 \\ 21 \quad 23 \\ 32 \quad 34 \\ 43 \quad 45 \\ 54 \quad 56 \\ 65 \quad 67 \\ 76 \quad 78 \\ 87 \quad 89 \\ 98 \end{array} \right\}$$

Как нетрудно проверить,

$$\mathbf{r} \circ \mathbf{r} = \left\{ \begin{array}{l} 00 \quad 02 \\ 11 \quad 13 \\ 20 \quad 22 \quad 24 \\ 31 \quad 33 \quad 35 \\ 42 \quad 44 \quad 46 \\ 53 \quad 55 \quad 57 \\ 64 \quad 66 \quad 68 \\ 75 \quad 77 \quad 79 \\ 86 \quad 88 \\ 97 \quad 99 \end{array} \right\}$$

Отношение r непосредственного соседства не является транзитивным: если a непосредственный сосед b и b непосредственный сосед c, то a может не быть непосредственным соседом c.

6.6. Отношения эквивалентности

Часто встречаются отношения, обладающие всеми указанными свойствами: рефлексивностью, симметричностью и транзитивностью. Такие отношения называются отношениями *эквивалентности.*

Определение отношения эквивалентности. *Каждое рефлексивное, симметричное и транзитивное отношение для множества A называется отношением эквивалентности для A.*

Пример 1. Отношение равенства — диагональ **e** квадрата $A \times A$ — является отношением эквивалентности для множества A.

Пример 2. Отношение четности

$$\mathbf{r} = \left\{\begin{matrix} 00 & & 02 & & 04 & & 06 & & 08 & \\ & 11 & & 13 & & 15 & & 17 & & 19 \\ 20 & & 22 & & 24 & & 26 & & 28 & \\ & 31 & & 33 & & 35 & & 37 & & 39 \\ 40 & & 42 & & 44 & & 46 & & 48 & \\ & 51 & & 53 & & 55 & & 57 & & 59 \\ 60 & & 62 & & 64 & & 66 & & 68 & \\ & 71 & & 73 & & 75 & & 77 & & 79 \\ 80 & & 82 & & 84 & & 86 & & 88 & \\ & 91 & & 93 & & 95 & & 97 & & 99 \end{matrix}\right\}$$

Это отношение является отношением эквивалентности для множества $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ цифр. Отношение «a имеет одинаковую четность с b» означает, что упорядоченная пара (a, b), образованная цифрами a и b, принадлежит отношению r.

7.6. Отношение порядка

Среди несимметричных отношений выделяются антисимметричные.

Отношение r для множества A называется *антисимметричным*, когда r обладает свойством: если элемент a связан отношением r с элементом b и, обратно, элемент b связан отношением r с элементом a, то элемент a равен элементу b. (Т. е. пересечение $\mathbf{r} \cap \mathbf{r}^{-1}$ отношения r с обратным отношением $\mathbf{r}^{-1}$ содержится в диагонали e квадрата $A \times A$.)

Пример 1. Обычное отношение порядка для цифр антисимметрично: для каждых цифр a и b, если $a \leqslant b$ и $b \leqslant a$, то $a = b$.

Пример 2. Отношение включения для класса $\mathcal{P}$ всех частей множества U антисимметрично: для каждых частей A и B множества U, если $A \subseteq B$ и $B \subseteq A$, то $A = B$.

Так же часто, как и отношения эквивалентности, встречаются рефлексивные, антисимметричные и транзитивные отношения.

Определение отношения порядка. *Каждое рефлексивное, антисимметричное и транзитивное отношение для множества A называется отношением порядка для A.*

Пример 1. Обычное отношение порядка для цифр рефлексивно, антисимметрично и транзитивно.

Пример 2. Отношение включения для класса $\mathcal{P}$ всех частей множества U является отношением порядка.

8.6. Отображения

Все отношения **r** между множествами A и B можно разделить на однозначные и многозначные.

Отношение **r** между множеством A и множеством B, при котором образ $\mathbf{r}(a)$ каждого элемента $a \in A$ состоит не более, чем из одного элемента множества B, называется *однозначным*. Неоднозначные отношения называются *многозначными*. При многозначном отношении образ некоторого элемента состоит из двух или более элементов.

Пример 1. Отношение равенства для цифр является однозначным отношением.

Пример 2. Отношение порядка для цифр является многозначным отношением.

Среди однозначных отношений **r** между множеством A и множеством B выделяются отношения, для которых множество A является областью определения. Это позволяет не рассматривать элементы a множества A, образы $\mathbf{r}(a)$ которых являются пустыми множествами. Такие отношения играют фундаментальную роль в математике.

Определение отображения. *Однозначное отношение f между множеством A и множеством B, область определения которого совпадает с множеством A, называется отображением множества A в множество B.*

Понятие отображения связано с интуитивно ясным понятием образа. В жизни на каждом шагу приходится людей, животных или вещи заменять некоторыми их образами: фотографиями, рисунками, описаниями, звуками.

Фраза *отображение f множества A в множество B* записывается символом

$$f: A \longrightarrow B.$$

Вместо «отображение f множества A в множество B» говорят также «функция f, определенная на множестве A и принимающая значения в множестве B». Вместо «*образ* $f(x)$ при отображении f элемента x» говорят также «*значение* $f(x)$ функции f для элемента x». Вместо символа $f: A \longrightarrow B$ используется также символ

$$f: x \longrightarrow f(x) \ (x \in A, \ f(x) \in B).$$

Пример 1. Отношение $f=e$ равенства для цифр является отображением множества $A=\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ в множество $B=A$. Каждая цифра x отображается при этом в цифру $f(x)=x$:

$$f(0)=0, f(1)=1, \ldots, f(9)=9.$$

Такое отображение называется *тождественным*.

Пример 2. Вещественная *квадратичная* функция f отображает множество $A=R$ всех вещественных чисел в множество $B=R_+$ всех положительных чисел. Каждое вещественное число x отображается при этом в положительное число $f(x)=x^2$. В частности,

$$f(0)=0, f(-1)=1, f(1)=1, f(\sqrt{2})=2.$$

Пример 3. Функция *число единиц среди элементов пары* отображает множество $A=\{00, 01, 10, 11\}$, составленное из упорядоченных пар 00, 01, 10, 11 в множество $B=\{0\ 1, 2\}$, составленное из натуральных чисел 0, 1, 2.

Эту функцию удобно определить таблицей

x	0 0	0 1	1 0	1 1
$f(x)$	0	1	1	2

Пример 4. Функция *сумма* отображает множество $A=\{(-1, -1), (-1, +1), (+1, -1), (+1, +1)$ в множество $B=\{-2, 0, +2\}$. Эту функцию удобно определить таблицей

x	$(-1, -1)$	$(-1, +1)$	$(+1, -1)$	$(+1, +1)$
$f(x)$	-2	0	0	2

Пример 5. Следующая таблица определяет функцию f на множестве $A=\{00, 01, 10, 11\}$ со значениями в множестве $B=[0, 1]$ всех положительных чисел, меньших единицы:

x	00	01	10	11
$f(x)$	$\frac{1}{4}\cdot\frac{1}{4}$	$\frac{1}{4}\cdot\frac{3}{4}$	$\frac{3}{4}\cdot\frac{1}{4}$	$\frac{3}{4}\cdot\frac{3}{4}$

Если область значений функции $f:A\to B$ совпадает с множеством B, то говорят, что f отображает множество A *на* множество B (примеры 1—4). Если отношение f^{-1}, обратное функции $f:A\to B$, однозначно, то функцию f называют *взаимно однозначной*. Взаимно однозначное отображение множества A на множество B называют также *изоморфизмом* множества A на

множество B или *подстановкой* вместо A множества B. Изоморфизм множества A на множество A (*на себя*) называется *автоморфизмом* или *перестановкой* множества A.

Пример 6. Следующие таблицы определяют перестановки f, g, h множества A цифр:

x	0 1 2 3 4 5 6 7 8 9
$f(x)$	0 1 2 3 4 5 6 7 8 9

;

x	0 1 2 3 4 5 6 7 8 9
$g(x)$	1 0 3 2 5 4 7 6 9 8

;

x	0 1 2 3 4 5 6 7 8 9
$h(x)$	9 8 7 6 5 4 3 2 1 0

.

Пример 7. Следующая таблица **определяет изоморфизм** φ множества $A=\{a, b, c, d, e, f, \ldots, z\}$ **букв латинского алфавита** на множество $B=\{1, 2, 3, 4, 5, 6, \ldots, 26\}$:

$\varphi(x)$	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26

Изоморфизм φ определяет нумерацию букв латинского алфавита.

Пример 8. Разобьем множество $\{00, 01, 10, 11\}$ всех упорядоченных пар, составленных из элементов 0 и 1, на части по числу единиц среди элементов пары:

$$f(0)=\{00\},\ f(1)=\{01, 10\},\ f(2)=\{11\}.$$

Тем самым определен изоморфизм f множества $A=\{0, 1, 2\}$ на класс

$$B=\{\{00\}, \{01, 10\}, \{11\}\}.$$

Пример 9. Сужение f_+ **вещественной квадратичной функции** на множество R_+ **всех положительных чисел является** ***автоморфизмом*** множества R_+ ($f_+(x)=x^2$ для каждого $x\geqslant 0$; для $x<0$ значение $f_+(x)$ не определено). Обратное для f_+ отношение f_+^{-1} является *арифметическим* квадратным корнем: $f_+^{-1}(y)=+\sqrt{y}$ для каждого числа $y>0$.

Пример 10. Сужение f_- вещественной квадратичной функции на множество R_- всех отрицательных чисел является *изоморфизмом* множества R_- на множество R_+ ($f_-(x)=x^2$ для каждого $x\leqslant 0$; для $x>0$ значение $f_-(x)$ не определено). Обратное для f_- отношение f_-^{-1} является *алгебраическим* квадратным корнем: $f_-^{-1}(y)=-\sqrt{y}$ для каждого числа $y\geqslant 0$.

Замечание. Вещественная квадратическая функция f **не является** изоморфизмом множества R всех вещественных чисел на множество R_+ всех положительных чисел, так как f не является взаимно однозначным отображением множества R на множество R_+. Отношение f^{-1}, обратное вещественной квадратичной функции f, многозначно: $f^{-1}(y) = \{+\sqrt{y}, -\sqrt{y}\}$, при этом $-\sqrt{y} \neq +\sqrt{y}$ для каждого числа $y > 0$.

§ 7. ПРИНЦИП ИНДУКЦИИ

Принцип индукции можно принять в качестве аксиомы. Этот принцип представляется интуитивно ясным и выражает одно из основных свойств упорядоченного множества

$$N = \{0, 1, 2, \ldots, n, n+1, \ldots\}$$

натуральных чисел.

1.7. Формулировка принципа индукции

Рассмотрим произвольную часть M множества N. Принцип индукции можно сформулировать следующим образом.

Если: 1) *число 0 принадлежит множеству M*,

2) *для каждого натурального числа n из того, что n принадлежит множеству M, следует, что число $n+1$ принадлежит M*,

то: *множество M совпадает с N*.

Используя сокращенную логическую символику, можно записать принцип индукции в форме, которая легко запоминается.

Принцип индукции. *Если*: 1) $0 \in M$,

2) $n \in M \Rightarrow n+1 \in M$,

то: $M = N$.

Стрелка в условии 2 заменяет оборот «*из ... следует ...*».

Используя интуицию, можно пояснить принцип индукции следующим рассуждением. По условию 1 число 0 принадлежит множеству M. По условию 2 отсюда следует, что число 1 принадлежит M. Значит, по этому же условию, число 2 принадлежит M. И так далее. Все натуральные числа принадлежат множеству M.

2.7. Формула суммы арифметической прогрессии

Докажем, используя принцип индукции, формулу для суммы n первых членов арифметической прогрессии с k-м членом $a+bk$:

$$(1) \qquad \sum_{0 \leqslant k < n} (a + bk) = an + b\frac{(n-1)n}{2}$$

для произвольных чисел a, b и натурального числа n.

Для того, чтобы доказать равенство (1), достаточно доказать равенство

$$\sum_{0 \leqslant k < n} k = \frac{(n-1)\,n}{2} \tag{2}$$

для каждого натурального числа n.

Если $n=0$, то суммы в левых частях равенств (1) и (2) по определению равны 0.

Рассмотрим множество M всех натуральных чисел m, для которых верно равенство

$$\sum_{0 \leqslant k < m} k = \frac{(m-1)\,m}{2}. \tag{3}$$

1. Число 0 принадлежит множеству M ($0 \in M$), так как при $m=0$ равенство (3) эквивалентно равенству $0=0$.

2. Для каждого натурального числа n из того, что n принадлежит множеству M, следует, что число $n+1$ принадлежит M ($n \in M \Rightarrow n+1 \in M$).

В самом деле, для каждого натурального числа n верно равенство

$$\sum_{0 \leqslant k < n+1} k = \sum_{0 \leqslant k < n} k + n. \tag{4}$$

Если число n принадлежит множеству M, то для него верно равенство (2) и из равенства (2) и (4) следуют равенства

$$\sum_{0 \leqslant k < n+1} k = \frac{(n-1)\,n}{2} + n = \frac{((n+1)-1)\,(n+1)}{2}.$$

Из этих равенств вытекает, что число $n+1$ принадлежит M.

Таким образом, условия 1 и 2 принципа индукции выполнены и, следовательно, множество M совпадает с множеством N всех натуральных чисел. Значит, равенство (2) верно для каждого натурального числа n.

3.7. «Теорема» о равенстве всех натуральных чисел

Для того, чтобы подчеркнуть важность тщательной проверки каждого из условий 1 и 2 принципа индукции, рассмотрим известный софизм.

«Теорема». *Все натуральные числа равны.*

«Доказательство» 1. Достаточно доказать, что каждое натуральное число l больше каждого натурального числа k или равно ему: $l \geqslant k$. Отсюда будет следовать, что для каждых натуральных чисел n и m верны неравенства $n \geqslant m$, $m \geqslant n$ и, значит, равенство $n=m$.

Рассмотрим множество M всех натуральных чисел m, больших или равных k.

1. Если $k=0$, то $0 \in M$.

2. Для каждого натурального числа n из $n \in M$ следует $n+1 \in M$: если $n \geqslant k$, то, тем более, $n+1 \geqslant k$.

По принципу индукции отсюда «следует», что $M=N$ и все натуральные числа больше числа k или равны ему.

Замечание. Условие 1 принципа индукции выполнено для $k=0$. В этом случае верно и высказанное утверждение о числе k. Если $k>0$, то число 0 не принадлежит множеству M, условие 1 принципа индукции не выполнено и применять этот принцип нельзя.

«Доказательство» 2. Определим множество M следующим образом: произвольное натуральное число m принадлежит множеству M, если каждые $m+1$ натуральных чисел равны, и не принадлежит M в противном случае.

1. Ясно, что $0 \in M$, так как каждое натуральное число равно самому себе.

2. Для каждого натурального числа $n>0$ из $n \in M$ следует $n+1 \in M$. Действительно, предположим, что каждые $n+1>1$ натуральных чисел равны. Рассмотрим произвольные $n+2>2$ натуральные числа $k_1, k_2, \ldots, k_{n+1}, k_{n+2}$. По предположению, числа $k_1, k_2, \ldots, k_{n+1}$ равны и числа $k_2, \ldots, k_{n+1}, k_{n+2}$ равны. Числа $k_2, \ldots, k_{n+1}$ принадлежат каждому из этих множеств.

Поэтому все рассматриваемые $n+2$ числа $k_1, k_2, \ldots, k_{n+1}, k_{n+2}$ равны.

По принципу индукции отсюда «следует», что $M=N$ и каждые $n+1$ натуральных чисел равны между собой.

Замечание. Условие 2 принципа индукции выполнено для $n>0$. Доказано, что если каждые два натуральных числа равны, то и все натуральные числа равны. Если $n=0$, то из $n \in M$ не следует $n+1 \in M$: из того, что каждое натуральное число равно самому себе, не следует, что каждые два натуральных числа равны. Условие 2 принципа индукции не выполнено, и применять этот принцип нельзя.

§ 8*. СЕМЕЙСТВА

Термин *семейство* эквивалентен термину *отображение*. Отображение f множества A в множество B называют *семейством элементов* множества B, имеющим *множество индексов* A. При этом вместо

$$f: x \to f(x) \quad (x \in A)$$

пишут

$$f = (f(x))_{x \in A}.$$

Вместо $f(x)$ пишут также f_x. Если из текста ясно, какие рассматриваются множество A и множество B, то семейство f обозначают символом $(f(x))_x$ или $(f_x)_x$.

1.8. Примеры

Рассмотрим несколько примеров семейств.

Пример 1. Каждое отображение f множества $A = N = \{0, 1, 2, \ldots\}$ всех натуральных чисел в множество B называется *последовательностью* элементов множества B и обозначается символом $(f_n)_{n \in N}$. Для обозначения последовательности f употребляется также символ

$$f_0, \ f_1, \ \ldots, \ f_n, \ldots$$

Пример 2. Каждое отображение f конечного множества A в множество B называется ***конечным семейством*** элементов множества B. В частности, конечными семействами являются отображения примеров 1, 3—8 § 6.

Замечание. Конечность семейства $f = (f(x))_{x \in A}$ означает конечность множества индексов A этого семейства, а не конечность его множества значений B. Например, последовательность f со значением $f(n) = (-1)^n \in \{-1, 1\}$ для каждого номера n не является конечным семейством, хотя его множество значений $B = \{-1, 1\}$ конечно.

Пример 3. Конечные семейства элементов множества B, имеющие множество индексов $A = \{1, 2, \ldots, n\}$, составленное из номеров $1, 2, \ldots, n$, называются ***строками*** длины n, составленными из элементов множества B, и обозначаются символом $(b_1, b_2, \ldots \ldots, b_n)$ или $b_1 b_2 \ldots b_n$. При этом говорят, что образ b_j номера j является *j-м элементом* строки $b_1 b_2 \ldots b_n$ или что он расположен на *j-м месте* в этой строке.

Иногда множество B называют ***алфавитом***, элементы множества B — буквами и вместо «*строка* длины n, составленная из элементов множества B», говорят «*слово* длины n, составленное из букв алфавита B».

В частности, *life*, *time*, *love*, *abcd*, *efgh* являются словами длины 4, составленными из букв латинского алфавита. Строки 00, 01, 10, 11 составляют множество всех слов длины 2, составленных из букв алфавита $B = \{0, 1\}$.

2.8. Сумма и произведение конечного семейства чисел

Часто приходится складывать или перемножать несколько чисел. Формально дело сводится к определению суммы и произведения конечного семейства чисел. Они определяются с помощью принципа индукции.

1. Для семейства $(f(x))_{x \in A}$ чисел, имеющего пустое множество индексов $A = O$, определяем сумму и произведение соответственно равенствами

$$\sum_{x \in O} f(x) = 0, \quad \prod_{x \in O} f(x) = 1.$$

Так определенные сумма и произведение коммутативны и ассоциативны.

2. Для каждого номера n, если для каждого семейства $(f(x))_{x\in A}$ чисел, имеющего множество индексов A, составленное из n элементов, определены коммутативные и ассоциативные сумма и произведение $\sum\limits_{x\in A} f(x)$, $\prod\limits_{x\in A} f(x)$, определяем сумму и произведение для каждого семейства $(f(x))_{x\in B}$ чисел, имеющего множество индексов B, составленное из $n+1$ элементов, соответственно равенствами

$$\sum_{x\in B} f(x) = \sum_{x\in A} f(x) + f(b),$$

$$\prod_{x\in B} f(x) = \prod_{x\in A} f(x)\cdot f(b).$$

В этих равенствах b обозначает произвольный элемент множества B и $A=B-\{b\}$. Из коммутативности и ассоциативности сложения и умножения, а также суммы и произведения произвольных n чисел вытекают однозначная определенность, коммутативность и ассоциативность суммы и произведения рассматриваемых произвольных $n+1$ чисел.

По принципу индукции из сказанного следует, что для каждого конечного семейства чисел определены коммутативные и ассоциативные сумма и произведение.

Сумма и произведение семейства чисел $(f(x))_{x\in A}$ обозначаются соответственно символами

$$\sum_{x\in A} f(x), \quad \prod_{x\in A} f(x).$$

Если множество A индексов составлено из номеров $0, 1, \ldots, m$, то используют также символы

$$\sum_{0\leqslant x\leqslant m} f(x) = \sum_{x=0}^{m} f(x) = f(0) + f(1) + \ldots + f(m),$$

$$\prod_{0\leqslant x\leqslant m} f(x) = \prod_{x=0}^{m} f(x) = f(0)\cdot f(1)\cdot \ldots \cdot f(m).$$

Такие символы для сумм были использованы в примерах 1 и 2 из § 7.

3.8. Примеры

Рассмотрим несколько примеров.

Пример 1. Если $A=\{0, 1, 2, 3\}$ и $f(x)=x$, то

$$\sum_{x\in A} f(x) = \sum_{0\leqslant x\leqslant 3} f(x) = f(0) + f(1) + f(2) + f(3) =$$

$$= 0 + 1 + 2 + 3 = 6,$$

$$\prod_{x\in A} f(x) = \prod_{0\leqslant x\leqslant 3} f(x) = f(0)\cdot f(1)\cdot f(2)\cdot f(3) = 0\cdot 1\cdot 2\cdot 3 = 0.$$

Пример 2. Если $A = \{(-1, -1), (-1, +1), (+1, -1), (+1, +1)\}$ и $f(-1, -1) = -2$, $f(-1, +1) = 0$, $f(+1, -1) = 0$, $f(+1, +1) = 2$, то

$$\sum_{x \in A} f(x) = f(-1, -1) + f(-1, +1) + f(+1, -1) + f(+1, +1) =$$
$$= -2 + 0 + 0 + 2 = 0.$$

Пример 3. Если $A = \{00, 01, 10, 11\}$ и $f(00) = \frac{1}{4} \cdot \frac{1}{4}$, $f(01) = \frac{1}{4} \cdot \frac{3}{4}$, $f(10) = \frac{3}{4} \cdot \frac{1}{4}$, $f(11) = \frac{3}{4} \cdot \frac{3}{4}$, то

$$\sum_{x \in A} f(x) = f(00) + f(01) + f(10) + f(11) = 1.$$

Пример 4. Коммутативность суммы означает, что для каждой перестановки p конечного множества индексов A равны сумма семейства $(f(x))_{x \in A}$ и сумма семейства $(f(p(x)))_{x \in A}$, полученного из него с помощью перестановки p:

$$\sum_{x \in A} f(x) = \sum_{x \in A} f(p(x)).$$

В частности, если множество A состоит из элементов 1 и 2, то дело сводится к коммутативности сложения:

$$f(1) + f(2) = f(2) + f(1).$$

Если A и f из примера 3, а перестановка p определяется таблицей

x	00	01	10	11
$p(x)$	11	10	01	00

то из коммутативности суммы следует равенство

$$\sum_{x \in A} f(x) = f(00) + f(01) + f(10) + f(11) =$$
$$= f(11) + f(10) + f(01) + f(00) = \sum_{x \in A} f(p(x)).$$

Пример 5. Ассоциативность суммы означает, что для каждого конечного семейства $(A_k)_{k \in K}$ непустых попарно не пересекающихся частей A_k конечного множества A, каждый элемент которого принадлежит некоторой из этих частей, равны сумма семейства $(f(x))_{x \in A}$ и сумма семейства $(g(k)_{k \in K}$, элементами которого являются суммы $g(k) = \sum_{x \in A_k} f(x)$ семейств $(f(x))_{x \in A_k}$:

$$\sum_{x \in A} f(x) = \sum_{k \in K} \sum_{x \in A_k} f(x).$$

В частности, если множество A состоит из элементов 1, 2 и 3, то дело сводится к ассоциативности сложения

$$f(1)+f(2)+f(3)=(f(1)+f(2))+f(3)=f(1)+(f(2)+f(3)).$$

Пример 6. Рассмотрим множество $A=\{000, 001, 010, 011, 100, 101, 110, 111\}$, образованное всеми строками длины 3, составленными из элементов 0 и 1. Рассмотрим множество $K=\{0; 1, 2, 3\}$ и семейство $(A_k)_{k\in K}$ непустых попарно не пересекающихся частей $A_0=\{000\}$, $A_1=\{001, 010, 100\}$, $A_2=\{011, 101, 110\}$, $A_3=\{111\}$ множества A. Каждый элемент множества A принадлежит некоторой из этих частей. Обозначим символом $s(x)$ число единиц в каждой строке x из множества A и рассмотрим число a положительное и меньшее единицы: $0\leqslant a\leqslant 1$. Определим семейство $(f(x))_{x\in A}$ равенствами

$$f(x)=a^{s(x)}(1-a)^{3-s(x)}.$$

В частности,

$$f(000)=(1-a)^3,\ f(001)=a(1-a)^2,$$
$$f(011)=a^2(1-a),\ f(111)=a^3.$$

В этом случае из ассоциативности сумм следуют равенства

$$\sum_{x\in A} f(x)=\sum_{k\in K}\sum_{x\in A_k} f(x)=\sum_{x\in A_0} f(x)+\sum_{x\in A_1} f(x)+\sum_{x\in A_2} f(x)+$$
$$+\sum_{x\in A_3} f(x)=f(000)+(f(001)+f(010)+f(100))+(f(011)+$$
$$+f(101)+f(110))+f(111)=(1-a)^3+3a(1-a)^2+3a^2(1-a)+$$
$$+a^3=((1-a)+a)^3=1.$$

4.8. Объединение и пересечение семейства множеств

По аналогии с определениями объединения и пересечения двух множеств определяются объединение и пересечение семейства множеств.

Рассмотрим семейство $(F(x))_{x\in A}$, элементы $F(x)$ которого являются частями множества U и которое имеет непустое множество индексов A. Объединением семейства $(F(x))_{x\in A}$ называется множество $\bigcup_{x\in A} F(x)$, образованное всеми элементами множества U, которые принадлежат *хотя бы одному* из множеств $F(x)$. Пересечением семейства $(F(x))_{x\in A}$ называется множество $\bigcap_{x\in A} F(x)$, образованное всеми элементами множества U, которые принадлежат *каждому* из множеств $F(x)$.

Кроме того, объединение и пересечение семейства $(F(x))_{x\in A}$ частей множества $F(x)$, которое имеет пустое множество индексов A, определяются соответственно равенствами

$$\bigcup_{x\in O} F(x)=O,\quad \bigcap_{x\in O} F(x)=U.$$

Замечание. Данное определение пересечения семейства множеств, имеющего пустое множество индексов, объясняется желанием сделать верными и для этого случая равенства

$$\Big(\bigcup_{x\in A} F(x)\Big)' = \bigcap_{x\in A} F'(x), \quad \Big(\bigcap_{x\in A} F(x)\Big)' = \bigcup_{x\in A} F'(x),$$

обобщающие теорему де Моргана.

Если множество индексов A конечно и составлено из номеров 0, 1, 2, ..., m, то для обозначения объединения и пересечения семейства $(F(x))_{x\in A}$ используют также символы:

$$\bigcup_{0\leqslant x\leqslant m} F(x) = \bigcup_{x=0}^{m} F(x) = F(0) \cup F(1) \cup \ldots \cup F(m),$$

$$\bigcap_{0\leqslant x\leqslant m} F(x) = \bigcap_{x=0}^{m} F(x) = F(0) \cap F(1) \cap \ldots \cap F(m).$$

Если множество индексов A состоит из двух элементов, то дело сводится к объединению и пересечению двух множеств, и используются обычные обозначения.

Непосредственно из определений следует, что объединение и пересечение семейства множеств коммутативны и ассоциативны в том же смысле, в котором коммутативны и ассоциативны сумма и произведение семейства чисел: перестановка множеств и их группировка не изменяют объединения и пересечения семейства.

Пример. Если $U=\{000, 001, 010, 011, 100, 101, 110, 111\}$, $A=\{0, 1, 2, 3\}$, $F(0)=\{000\}$, $F_1=\{001, 010, 100\}$, $F_2=\{011, 101, 110\}$, $F_3=\{111\}$, то, как нетрудно проверить,

$$\bigcup_{x\in A} F(x) = \bigcup_{0\leqslant x\leqslant 3} F(x) = F(0) \cup F(1) \cup F(2) \cup F(3) = U,$$

$$\bigcap_{x\in A} F(x) = \bigcap_{0\leqslant x\leqslant 3} F(x) = F(0) \cap F(1) F(2) \cap F(3) = 0.$$

Глава 2

ЭЛЕМЕНТЫ КОМБИНАТОРИКИ

Комбинаторика изучает способы подсчета числа элементов различных конечных множеств. Важную роль в комбинаторике играют правила сложения и умножения. С их помощью решаются, в частности, классические задачи о числе перестановок, выборок и размещений.

При первом чтении рекомендуется ограничиться параграфами 1—5 этой главы.

Определения встречающихся терминов можно найти в главе 1.

Задачи к главе 2 собраны в главе 1 части II.

§ 1. ПРАВИЛА СЛОЖЕНИЯ И УМНОЖЕНИЯ

Правило сложения для двух множеств выражает основное свойство числа элементов конечного множества. Это правило можно принять в качестве аксиомы. Интуитивно оно представляется очевидным.

1.1. Правило сложения

Условимся число элементов конечного множества A обозначать символом $n(A)$.

Пример 1. Число элементов пустого множества O равно нулю: $n(O)=0$.

Пример 2. Число элементов множества $\{a\}$, образованного единственным элементом a, равно единице: $n(\{a\})=1$.

Пример 3. Число элементов множества $F=\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, составленного из цифр, равно десяти: $n(F)=10$.

Правило сложения. *Число $n(A+B)$ элементов суммы $A+B$ не пересекающихся конечных множеств A и B равно сумме $n(A)+n(B)$ чисел $n(A)$ и $n(B)$ элементов этих множеств.*

Правило сложения можно записать коротко следующим образом:

$$n(A+B)=n(A)+n(B) \quad (AB=O).$$

Пример 1.

$$n(\{1, 2, 3, 4, 5, 6\})=n(\{2, 4, 6\})+n(\{1, 3, 5\}).$$

Пример 2.

$$n(\{a, b, c, \ldots, y, z\})=n(\{a, e, i, o, u\})+n(\{b, c, d, \ldots, y, z\}).$$

Пример 3.

$$n(\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}\neq n(\{0, 1, 2, 3, 4, 5\})+n(\{5, 6, 7, 8, 9\}).$$

Замечание. Пример 3 показывает, что условие $AB=O$ об отсутствии у множеств A и B общих элементов существенно для правила сложения.

2.1*. Общее правило сложения

Используя принцип индукции, легко вывести из правила сложения общее правило сложения для конечного семейства множеств.

Рассмотрим семейство $(F(x)_{x\in A}$, элементы $F(x)$ которого являются попарно не пересекающимися конечными частями множества U и которое имеет конечное множество индексов A. Условимся объединение множеств этого семейства называть также его суммой и обозначать символом $\sum\limits_{x\in A} F(x)$.

Пример 1.

$$U = \{1, 2, 3, 4, 5, 6\},$$
$$F(0) = \{2, 4, 6\},\ F(1) = \{1, 3, 5\},$$
$$A = \{0, 1\},$$
$$\sum_{x \in A} F(x) = F(0) + F(1) = U.$$

Пример 2.

$$U = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\},$$
$$F(0) = \{0, 1, 2\},\ F(3) = \{3, 4, 5\},\ F(6) = \{6, 7, 8\},$$
$$A = \{0, 3, 6\},$$
$$\sum_{x \in A} F(x) = F(0) + F(3) + F(6) = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}.$$

Пример 3.

$$U = \{000, 001, 010, 100, 011, 101, 110, 111\},$$
$$F(0) = \{000\},\quad F(1) = \{001,\ 010,\ 100\},\quad F(2) = \{011,\ 101,\ 110\},$$
$$F(3) = \{111\},\qquad A = \{0, 1, 2, 3\},$$
$$\sum_{x \in A} F(x) = F(0) + F(1) + F(2) + F(3) = U.$$

Общее правило сложения. *Число $n(\Sigma F(x))$ элементов суммы $\Sigma F(x)$ попарно не пересекающихся конечных множеств $F(x)$ равно сумме $\Sigma n(F(x))$ чисел $n(F(x))$ элементов этих множеств.*

— Рассмотрим множество M всех номеров m таких, что общее правило сложения верно для всех семейств $(F(x))_{x \in A}$ попарно не пересекающихся конечных частей $F(x)$ множества U, имеющих множество индексов A, составленное из m элементов.

1. $0 \in M$. Это вытекает из определений. В самом деле, если множество A имеет 0 элементов (т. е. является пустым множеством O), то, по определению, сумма $\sum_{x \in O} F(x)$ пустого семейства множеств $(F(x))_{x \in O}$ равна пустому множеству O, а сумма $\sum_{x \in O} n(F(x))$ пустого семейства чисел $(n(F(x)))_{x \in O}$ — числу 0. И следовательно,

$$n\left(\sum_{x \in O} F(x)\right) = n(O) = 0 = \sum_{x \in O} n(F(x)).$$

2. Для каждого натурального числа n, если $n \in M$, то $n+1 \in M$. В самом деле, рассмотрим семейство $(F(x))_{x \in B}$ попарно не пересекающихся конечных частей $F(x)$ множества U, имеющее множество индексов B, составленное из $n+1$ элементов. Из пра-

вила сложения вытекает, что для каждого элемента $b\in B$ и множества $A=B-\{b\}$ остальных элементов верны равенства

$$n\left(\sum_{x\in B} F(x)\right) = n\left(\sum_{x\in A} F(x) + F(b)\right) = n\left(\sum_{x\in A} F(x)\right) + n(F(b)).$$

Множество A состоит из n элементов. Поэтому, если $n\in M$, то

$$n\left(\sum_{x\in A} F(x)\right) = \sum_{x\in A} n(F(x)),$$

и, следовательно,

$$n\left(\sum_{x\in B} F(x)\right) = \sum_{x\in A} n(F(x) + n(F(b)) = \sum_{x\in B} n(F(x)),$$

т. е. $n+1\in M$.

Таким образом, условия 1 и 2 принципа индукции выполнены, значит, множество M совпадает с множеством N всех натуральных чисел. Общее правило сложения верно. +

Пример 1.

$$n(\{2, 4, 6\}+\{1, 3, 5\}) = n(\{2, 4, 6\}+n(\{1, 3, 5\}).$$

Пример 2.

$$n(\{000\}+\{001, 010, 100\}+\{011, 101, 110\}+\{111\}) =$$
$$= n(\{000\})+n(\{001, 010, 100\})+n(\{011, 101, 110\})+n(\{111\}).$$

Пример 3.

$$n(\{0, 1, 2\}+\{3, 4, 5\}+\{6, 7, 8\}) = n(\{0, 1, 2\})+n(\{3, 4, 5\})+$$
$$+n(\{6, 7, 8\}).$$

Замечание. Пример 1 повторяет пример 1 пункта 1.1. Это повторение поясняет тот факт, что правило сложения для двух множеств является частным случаем общего правила сложения.

3.1. Правило умножения

Из общего правила сложения выводится правило умножения.

Рассмотрим произвольные множества A и B. Каждые элементы $a\in A$ и $b\in B$ определяют упорядоченную пару ab. Множество всех таких пар образует декартово произведение $A\times B$ множеств A и B.

Пример 1.

$$A=B=\{0, 1, \ldots, 9\}, \qquad A\times B=\{00, 01, \ldots, 99\}.$$

Пример 2.

$$A=\{1, 2, \ldots, 8\},\ B=\{a, b, \ldots, h\},$$
$$A\times B=\{1a, 1b, \ldots, 8h\}.$$

Пример 3.

$$A=\{1, 2, \ldots, 8\},\ B=\{a, b, \ldots, h\},$$
$$B\times A=\{a1, a2, \ldots, h8\}.$$

Замечание. Примеры 2 и 3 показывают, что $A\times B\neq B\times A$.

Правило умножения. *Число $n(A\times B)$ элементов декартова произведения $A\times B$ конечных множеств A и B равно произведению $n(A)\times n(B)$ чисел $n(A)$ и $n(B)$ элементов этих множеств.*

— Для каждого элемента $x\in A$ обозначим символом $F(x)$ множество всех упорядоченных пар xy, составленных из элемента x и произвольного элемента $y\in B$:

$$F(x)=\{x\}\times B.$$

Эти множества попарно не пересекаются, и число элементов в каждом из них равно числу элементов множества B:

$$n(F(x))=n(B).$$

Сумма всех множеств $F(x)$ равна декартову произведению $A\times B$ множеств A и B:

$$\sum_{x\in A} F(x)=\sum_{x\in A}\{x\}\times B=A\times B.$$

По правилу сложения из сказанного следует, что

$$n(A\times B)=\sum_{x\in A} n(\{x\}\times B)=n(A)\times n(B).$$

Правило умножения доказано. +

Пример 1.

$$n(\{0, 1, \ldots, 9\}\times\{0, 1, \ldots, 9\})=n(\{0, 1, \ldots, 9\})\times n(\{0, 1, \ldots, 9\})=100.$$

Пример 2.

$$n(\{1, 2, \ldots, 8\}\times\{a, b, \ldots, h\})=$$
$$=n(\{1, 2, \ldots, 8\})\times n(\{a, b, \ldots, h\})=64.$$

Пример 3.

$$n(\{0, 1\}\times\{0, 1\})=n(\{0, 1\})\times n(\{0, 1\})=4.$$

4.1*. Общее правило умножения

Используя принцип индукции, легко вывести из правила умножения общее правило умножения для конечного семейства множеств.

Рассмотрим семейство $(F(x))_{x\in A}$, элементы $F(x)$ которого являются частями множества B и которое имеет множество ин-

дексов A. Рассмотрим также часть $\prod_{x \in A} F(x)$ множества $\mathscr{P}$ всех частей произведения $A \times B$, образованную всеми семействами $(f(x))_{x \in A}$, такими, что $f(x) \in F(x)$ для каждого $x \in A$. Условимся называть множество $\prod_{x \in A} F(x)$ декартовым произведением семейства $(F(x))_{x \in A}$. Определение декартова произведения семейства множеств можно записать равенством

$$\prod_{x \in A} F(x) = \{(f(x))_{x \in A} : f(x) \in F(x) \quad (x \in A)\}.$$

В частности, если $A = O$, то единственным элементом множества $\prod_{x \in O} F(x)$ является пустая часть множества $A \times B$.

Если $F(x) = B$ для каждого индекса $x \in A$, то декартово произведение семейства $(F(x))_{x \in A}$ называется декартовой степенью A множества B и обозначается символом B^A. Это множество образовано всеми отображениями множества A в множество B или, что эквивалентно, всеми семействами элементов множества B, имеющими множество индексов A.

Если множество индексов A составлено из номеров, то для обозначения декартовых произведений множеств используются символы, аналогичные обозначающим произведения чисел.

Пример 1. Если $A = \{1, 2, 3\}$, $B = \{0, 1\}$ и $F(1) = \{0\}$, $F(2) = \{1\}$, $F(3) = \{0, 1\}$, то

$$\prod_{x \in A} F(x) = F(1) \times F(2) \times F(3) = \{0\} \times \{1\} \times \{0, 1\} = \{010, 011\}.$$

Пример 2. Если $B = \{0, 1\}$, $A = \{1, 2, 3\}$ и $F(1) = F(2) = F(3) = B$, то

$$\prod_{x \in A} F(x) = B^A = \{0, 1\} \times \{0, 1\} \times \{0, 1\} =$$
$$= \{000, 001, 010, 011, 100, 101, 110, 111\}.$$

Пример 3. Если $B = R$ является множеством всех вещественных чисел, а $A = N$ — множеством всех натуральных чисел, то $B^A = R^N$ является множеством всех последовательностей вещественных чисел.

Рассмотрим семейство $(F(x))_{x \in A}$ конечных частей $F(x)$ множества B, имеющее конечное множество индексов A.

Общее правило умножения. *Число $n(\Pi F(x))$ элементов декартова произведения $\Pi F(x)$ конечных множеств $F(x)$ равно произведению $\Pi n(F(x))$ чисел $n(F(x))$ элементов этих множеств.*

— Рассмотрим множество M всех номеров m таких, что общее правило умножения верно для всех семейств $(F(x))_{x \in A}$ конечных частей $F(x)$ множества B, имеющих множество индексов A, составленное из m элементов.

1. $0 \in M$. В самом деле, если $n(A) = 0$ и, следовательно, $A = 0$, то единственным элементом произведения $\prod_{x \in O} F(x)$ является

пустое семейство. В то же время произведение $\prod_{x\in O} n(F(x))$, по определению, равно единице. Значит,

$$n\left(\prod_{x\in O} F(x)\right) = 1 = \prod_{x\in O} n(F(x)).$$

2. Для каждого натурального числа n, если $n\in M$, то $n+1\in M$. В самом деле, рассмотрим семейство $(F(x))_{x\in C}$ конечных частей $F(x)$ множества B, имеющее множество индексов C, составленное из $n+1$ элементов. Рассмотрим также произвольный элемент $c\in C$ и множество $A=C-\{c\}$ остальных элементов. Отображение

$$(f(x))_{x\in C} \longrightarrow ((f(x))_{x\in A}, f(c))$$

является изоморфизмом множества $\prod_{x\in C} F(x)$ на множество $\prod_{x\in A} F(x)\times F(c)$. Следовательно, число элементов в этих множествах одинаково:

$$n\left(\prod_{x\in C} F(x)\right) = n\left(\prod_{x\in A} F(x)\times F(c)\right).$$

Из правила умножения вытекает, что

$$n\left(\prod_{x\in A} F(x)\times F(c)\right) = n\left(\prod_{x\in A} F(x)\right)\times n(F(c)).$$

Множество A состоит из n элементов. Поэтому, если $n\in M$, то

$$n\left(\prod_{x\in A} F(x)\right) = \prod_{x\in A} n(F(x))$$

и, следовательно,

$$n\left(\prod_{x\in C} F(x)\right) = \prod_{x\in A} n(F(x))\times n(F(c)) = \prod_{x\in C} n(F(x)),$$

т. е. $n+1\in M$.

Таким образом, условия 1 и 2 принципа индукции выполнены и, значит, множество M совпадает с множеством N всех натуральных чисел. Общее правило умножения верно. +

Пример 1. Если $B=\{0, 1\}$, $A=\{1, 2, 3\}$ и $F(1)=\{0\}$, $F(2)=\{1\}$, $F(3)=\{0, 1\}$, то

$$n(\prod F(x)) = \prod n(F(x)) = n(F(1))\times n(F(2))\times n(F(3)) = 1\times 1\times 2 = 2.$$

Пример 2. Если $B=\{0, 1\}$, $A=\{1, 2, 3\}$ и $F(1)=F(2)=F(3)=B$, то

$$n(\prod F(x)) = n(B^A) = n(B)\times n(B)\times n(B) = n(B)^{n(A)} = 2^3 = 8.$$

Пример 3. Для каждых конечных множеств A и B верно равенство

$$n(B^A) = n(B)^{n(A)}.$$

Задачи к § 1 собраны в § 1 главы 1 части II.

§ 2. ПЕРЕСТАНОВКИ

Классической задачей комбинаторики является задача о числе перестановок, содержание которой можно выразить вопросом: *сколькими способами можно переставить n различных предметов, расположенных на n различных местах*?

1.2. Примеры

Рассмотрим несколько примеров.

Пример 1. *Сколькими способами можно переставить три монеты 1, 2, 3 копейки, расположенные соответственно на трех местах с номерами 1, 2, 3*?

Каждую перестановку монет 1, 2, 3 копейки на местах с номерами 1, 2, 3 можно описать перестановкой s множества $A=\{1, 2, 3\}$. В частности, если каждая монета остается на своем месте, то соответствующая перестановка является тождественной и описывается таблицами

x	1 2 3
$s(x)$	1 2 3

; $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$.

Если монеты 1 и 2 меняются местами, то соответствующая перестановка описывается таблицами

x	1 2 3
$s(x)$	2 1 3

; $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$.

Вообще, если каждая монета x переставляется на место $s(x)$, то соответствующая перестановка описывается таблицей, верхняя строка которой составлена из значений x, а нижняя — из расположенных под ними соответствующих значений $s(x)$:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Нетрудно непосредственно проверить, что эти таблицы описывают все возможные перестановки множества $A=\{1, 2, 3\}$.

Таким образом, монеты 1, 2, 3 копейки на местах с номерами 1, 2, 3 можно переставить

$$1\cdot2\cdot3=6$$

способами.

Пример 2. *Сколькими способами можно пересадить четырех гостей $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$, сидящих соответственно на четырех местах 1, 2, 3, 4*?

Каждую пересадку гостей $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$, $\mathscr{D}$ на местах 1, 2, 3, 4 можно описать подстановкой вместо $A = \{1, 2, 3, 4\}$ множества $B = \{\mathscr{A}, \mathscr{B}, \mathscr{C}, \mathscr{D}\}$. Если на каждое место x пересаживается гость $s(x)$, соответствующая подстановка описывается таблицей, верхняя строка которой составлена из значений x, а нижняя — из расположенных под ними соответствующих значений $s(x)$:

$$\begin{pmatrix}1&2&3&4\\ \mathscr{A}&\mathscr{B}&\mathscr{C}&\mathscr{D}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{B}&\mathscr{A}&\mathscr{C}&\mathscr{D}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{C}&\mathscr{B}&\mathscr{A}&\mathscr{D}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{B}&\mathscr{C}&\mathscr{D}&\mathscr{A}\end{pmatrix},$$

$$\begin{pmatrix}1&2&3&4\\ \mathscr{A}&\mathscr{C}&\mathscr{B}&\mathscr{D}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{C}&\mathscr{A}&\mathscr{B}&\mathscr{D}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{C}&\mathscr{B}&\mathscr{A}&\mathscr{D}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{C}&\mathscr{B}&\mathscr{D}&\mathscr{A}\end{pmatrix},$$

$$\begin{pmatrix}1&2&3&4\\ \mathscr{A}&\mathscr{D}&\mathscr{C}&\mathscr{B}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{D}&\mathscr{A}&\mathscr{C}&\mathscr{B}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{D}&\mathscr{C}&\mathscr{A}&\mathscr{B}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{D}&\mathscr{C}&\mathscr{B}&\mathscr{A}\end{pmatrix},$$

$$\begin{pmatrix}1&2&3&4\\ \mathscr{A}&\mathscr{B}&\mathscr{D}&\mathscr{C}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{B}&\mathscr{A}&\mathscr{D}&\mathscr{C}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{B}&\mathscr{D}&\mathscr{A}&\mathscr{C}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{B}&\mathscr{D}&\mathscr{C}&\mathscr{A}\end{pmatrix},$$

$$\begin{pmatrix}1&2&3&4\\ \mathscr{A}&\mathscr{D}&\mathscr{B}&\mathscr{C}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{D}&\mathscr{A}&\mathscr{B}&\mathscr{C}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{D}&\mathscr{B}&\mathscr{A}&\mathscr{C}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{D}&\mathscr{B}&\mathscr{C}&\mathscr{A}\end{pmatrix},$$

$$\begin{pmatrix}1&2&3&4\\ \mathscr{A}&\mathscr{C}&\mathscr{D}&\mathscr{B}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{C}&\mathscr{A}&\mathscr{D}&\mathscr{B}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{C}&\mathscr{D}&\mathscr{A}&\mathscr{B}\end{pmatrix}, \begin{pmatrix}1&2&3&4\\ \mathscr{C}&\mathscr{D}&\mathscr{B}&\mathscr{A}\end{pmatrix}.$$

Нетрудно убедиться в том, что эта таблица описывает все возможные подстановки вместо $A = \{1, 2, 3, 4\}$ множества $B = \{\mathscr{A}, \mathscr{B}, \mathscr{C}, \mathscr{D}\}$. Действительно, в букву $\mathscr{A}$ отображается либо номер 1, либо 2, либо 3, либо 4. В каждом из этих случаев в остающиеся три буквы $\mathscr{B}$, $\mathscr{C}$, $\mathscr{D}$ произвольным образом отображаются остающиеся три номера: либо 2, 3, 4, либо 1, 3, 4, либо 1, 2, 4, либо 1, 2, 3. Соответствующие подстановки описываются столбцами полученной таблицы подстановок.

Таким образом, гостей $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$, $\mathscr{D}$ на местах 1, 2, 3, 4. можно пересадить

$$6 \cdot 4 = 1 \cdot 2 \cdot 3 \cdot 4 = 24$$

способами.

Пример 3. *Сколькими способами можно переставить буквы в слове camel?*

Каждую перестановку букв в слове *camel* можно описать подстановкой вместо $A = \{1, 2, 3, 4, 5\}$ множества $B = \{c, a, m, e, l\}$. Если на каждое место x в слове ставится буква $s(x)$, то соответствующая подстановка описывается таблицей, верхняя строка которой составлена из значений x, а нижняя — из расположенных под ними соответствующих значений $s(x)$. Например,

$$\begin{pmatrix}1&2&3&4&5\\ c&a&m&e&l\end{pmatrix}, \begin{pmatrix}1&2&3&4&5\\ a&c&m&e&l\end{pmatrix}, \begin{pmatrix}1&2&3&4&5\\ c&m&a&e&l\end{pmatrix}, \begin{pmatrix}1&2&3&4&5\\ c&m&e&a&l\end{pmatrix}.$$

По аналогии с примером 2 подсчитаем число всех таких подстановок. В букву a отображается либо номер 1, либо 2, ..., либо 5. В каждом из этих случаев в остающиеся четыре буквы c, m, e, l произвольным образом отображаются остающиеся четыре номера: либо 2, 3, 4, 5, либо 1, 3, 4, 5,..., либо 1, 2, 3, 4. Из примера 2 ясно, что таких подстановок 24.

Таким образом, буквы в слове *camel* можно переставить

$$24\cdot 5=1\cdot 2\cdot 3\cdot 4\cdot 5=120$$

способами.

2.2. Формула для числа подстановок

Рассмотрим произвольное натуральное число n и множества A, B, состоящие каждое из n элементов. По определению, *подстановкой вместо A множества B* называется взаимно однозначное отображение A на B. Обозначим символом $S(A, B)$ множество всех подстановок вместо A множества B. Число этих подстановок, т. е. число элементов множества $S(A, B)$, определяется числом n элементов множеств A, B и обозначается символом $n!$ (эн-факториал).

В частности, если $n=0$ и множества A, B пустые, то множество $S(A, B)$ состоит из пустой подстановки и

$$0!=1.$$

Если $n=1$ и каждое из множеств A, B состоит из единственного элемента, то, очевидно, множество $S(A, B)$ состоит из единственной подстановки вместо элемента из A элемента множества B. Следовательно,

$$1!=1.$$

Если $n=2$ и каждое из множеств A, B состоит из двух элементов:

$$A=\{a, b\},\ B=\{c, d\},$$

то множество $S(A, B)$ состоит из двух подстановок:

$$S(A,B)=\left\{\begin{pmatrix} a & b \\ c & d\end{pmatrix}, \begin{pmatrix} a & b \\ d & c\end{pmatrix}\right\}.$$

Следовательно,

$$2!=1!\ 2=1\cdot 2=2.$$

Подсчет числа подстановок в примерах 1, 2, 3 показывает, что

$$3!=2!\cdot 3=1\cdot 2\cdot 3,\ 4!=3!\cdot 4=1\cdot 2\cdot 3\cdot 4,$$
$$5!=4!\cdot 5=1\cdot 2\cdot 3\cdot 4\cdot 5.$$

Эти равенства подсказывают общую формулу для числа подстановок. Используя общее правило сложения и принцип индук-

ции, нетрудно доказать, что для каждых натурального числа $n>0$, множества A из n элементов и множества B из n элементов число $n!$ всех подстановок вместо A множества B равно произведению $1\cdot 2\cdot \ldots \cdot n$ чисел $1, 2, \ldots, n$.

Для $n=0$ символом $1\cdot 2\cdot \ldots \cdot n$ будем обозначать число 1. При этом условии для каждого натурального числа n верна

Формула для числа подстановок:

$$n!=1\cdot 2\cdot \ldots \cdot n.$$

— Рассмотрим множество M всех натуральных чисел m, для которых $m!=1\cdot 2\cdot \ldots \cdot m$, т. е. таких, что для каждых множества A из m элементов и множества B из m элементов число всех подстановок вместо A множества B равно $1\cdot 2\cdot \ldots \cdot m$.

1. Число 0 принадлежит множеству M, так как $m!=1=$ $=1\cdot 2\cdot \ldots \cdot m$ при $m=0$.

2. Для каждого натурального числа n верно предложение — если n принадлежит M, то $n+1$ принадлежит M.

Действительно, рассмотрим произвсльные натуральное число n и множества A, B, состоящие из $n+1$ элементов. Рассмотрим также произвольный элемент b множества B. Для каждого элемента x множества A обозначим символом $F(x)$ множество всех подстановок вместо A множества B, при которых x отображается в b. Множества $F(x)$ попарно не пересекаются, и их сумма равна множеству $S(A, B)$ всех подстановок вместо A множества B. Это вытекает непосредственно из определений. Если подстановка s принадлежит множествам $F(x)$ и $F(y)$ для некоторых x и y из множества A, то из взаимной однозначности s и равенств $s(x)=b=s(y)$ следует равенство $x=y$. Значит, при каждых различных x и y множества $F(x)$ и $F(y)$ не пересекаются. Каждая подстановка s вместо A множества B отображает A на B. Следовательно, $s(x)=b$ для некоторого $x\in A$, s принадлежит множеству $F(x)$ и сумма всех множеств $F(x)$ равна множеству $S(A, B)$:

$$\Sigma F(x)=S(A, B). \quad (1)$$

Для каждого $x\in A$ множество $F(x)$ подстановок вместо A множества B, отображающих x в b, изоморфно множеству $S(A-\{x\}, B-\{b\})$ всех подстановок вместо $A-\{x\}$ множества $B-\{b\}$. Действительно, каждая подстановка $s\in F(x)$ определяет подстановку $t\in S(A-\{x\}, B-\{b\})$ со значением $t(y)=s(y)$ для каждого $y\in A-\{x\}$. Различные подстановки s из $F(x)$ определяют по этому правилу различные подстановки t, и каждая подстановка t из $S(A-\{x\}, B-\{b\})$ определяется некоторой подстановкой s. Следовательно, правило определяет изоморфизм (взаимно однозначное отображение) $F(x)$ на $S(A-\{x\}, B-\{b\})$. Значит, эти множества имеют одно и то же число элементов $n!$.

Если n принадлежит множеству M, то

$$n!=1\cdot 2\cdot \ldots \cdot n. \quad (2)$$

По правилу сложения число элементов суммы всех множеств $F(x)$ равно сумме чисел элементов этих множеств:

$$(3) \qquad n(\Sigma F(x)) = \Sigma n(F(x)).$$

Из равенств (1) — (3) следует, что

$$(n+1)! = n(\Sigma F(x)) = \Sigma n(F(x)) = n!(n+1) = 1 \cdot 2 \cdot \ldots \cdot n(n+1).$$

По принципу индукции из доказанного вытекает, что формула для числа подстановок верна. $+$

Пример 1. Рассмотрим множества $A = \{1, 2, 3\}$ и $B = \{a, b, c\}$. В этом случае

$$F(1) = \left\{\begin{pmatrix}1 & 2 & 3\\ b & a & c\end{pmatrix}, \begin{pmatrix}1 & 2 & 3\\ b & c & a\end{pmatrix}\right\}, \quad S(\{2,3\},\{a,c\}) = \left\{\begin{pmatrix}2 & 3\\ a & c\end{pmatrix}, \begin{pmatrix}2 & 3\\ c & a\end{pmatrix}\right\};$$

$$F(2) = \left\{\begin{pmatrix}1 & 2 & 3\\ a & b & c\end{pmatrix}, \begin{pmatrix}1 & 2 & 3\\ c & b & a\end{pmatrix}\right\}, \quad S(\{1,3\},\{a,c\}) = \left\{\begin{pmatrix}1 & 3\\ a & c\end{pmatrix}, \begin{pmatrix}1 & 3\\ c & a\end{pmatrix}\right\};$$

$$F(3) = \left\{\begin{pmatrix}1 & 2 & 3\\ a & c & b\end{pmatrix}, \begin{pmatrix}1 & 2 & 3\\ c & a & b\end{pmatrix}\right\}, \quad S(\{1,2\},\{a,c\} = \left\{\begin{pmatrix}1 & 2\\ a & c\end{pmatrix}, \begin{pmatrix}1 & 2\\ c & a\end{pmatrix}\right\}.$$

Первые подстановки s множеств F определяют подстановки t соответствующих множеств S, вторые — вторые подстановки из S.

Пример 2. В примере 1 из пункта 1.2 при $b = \mathscr{A}$ множества F описываются столбцами рассматриваемой там таблицы подстановок.

Из формулы для числа подстановок, в частности, следует, что число всех перестановок (т. е. взаимно однозначных отображений на себя) произвольного множества из n элементов равно $n! = 1 \cdot \ldots \cdot n$.

Формула для числа подстановок дает решение классической задачи о числе перестановок: *n различных предметов, расположенных на n различных местах, можно переставить $n! = 1 \cdot \ldots \cdot n$ способами.*

§ 3. ВЫБОРКИ

Классической задачей комбинаторики является задача о числе выборок, содержание которой можно выразить вопросом: *сколькими способами можно выбрать m из n различных предметов?*

1.3. Примеры

Рассмотрим несколько примеров.

Пример 1. *Сколькими способами можно выбрать две из трех монет 1, 2, 3 копейки?*

Каждую выборку $m = 2$ из $n = 3$ монет 1, 2, 3 копейки можно описать частью Y множества $B = \{1, 2, 3\}$, составленной из $m = 2$ элементов. Если выбираются монеты 1, 2 копейки, то $Y = \{1, 2\}$,

1, 3 копейки — $Y=\{1, 3\}$, 2, 3 копейки — $Y=\{2, 3\}$. Эти части составляют класс

$$\mathscr{P}_2(\{1, 2, 3\})=\{\{1, 2\}, \{1, 3\}, \{2, 3\}\}.$$

Нетрудно установить зависимость между числом $\binom{3}{2}$ множеств этого класса и числом 3! перестановок множества $B=\{1, 2, 3\}$. Каждые подстановки s вместо $A=\{1, 2\}$ множества Y и подстановка t вместо $B—A=\{3\}$ множества $B—Y$ определяют перестановку u множества $B=\{1, 2, 3\}$:

Y	s	t	u
$\{1,2\}$	$\begin{pmatrix}1 & 2\\1 & 2\end{pmatrix}$	$\begin{pmatrix}3\\3\end{pmatrix}$	$\begin{pmatrix}1 & 2 & 3\\1 & 2 & 3\end{pmatrix}$
	$\begin{pmatrix}1 & 2\\2 & 1\end{pmatrix}$	$\begin{pmatrix}3\\3\end{pmatrix}$	$\begin{pmatrix}1 & 2 & 3\\2 & 1 & 3\end{pmatrix}$
$\{1,3\}$	$\begin{pmatrix}1 & 2\\1 & 3\end{pmatrix}$	$\begin{pmatrix}3\\2\end{pmatrix}$	$\begin{pmatrix}1 & 2 & 3\\1 & 3 & 2\end{pmatrix}$
	$\begin{pmatrix}1 & 2\\3 & 1\end{pmatrix}$	$\begin{pmatrix}3\\2\end{pmatrix}$	$\begin{pmatrix}1 & 2 & 3\\3 & 1 & 2\end{pmatrix}$
$\{2,3\}$	$\begin{pmatrix}1 & 2\\2 & 3\end{pmatrix}$	$\begin{pmatrix}3\\1\end{pmatrix}$	$\begin{pmatrix}1 & 2 & 3\\2 & 3 & 1\end{pmatrix}$
	$\begin{pmatrix}1 & 2\\3 & 2\end{pmatrix}$	$\begin{pmatrix}3\\1\end{pmatrix}$	$\begin{pmatrix}1 & 2 & 3\\3 & 2 & 1\end{pmatrix}$

В этой таблице для каждого множества Y в соответствующих строках столбцов s и t расположены подстановки из множеств $S(A, Y)$ и $S(B—A, B—Y)$. Множества $F(Y)=S(A, Y)\times S(B—A, B—Y)$ пар подстановок s, t не пересекаются и в сумме образуют множество, число элементов которого равно числу всех перестановок множества B. В соответствии с правилом умножения число пар подстановок s, t в каждом из множеств $F(Y)$ равно произведению $2!(3—2)!$ числа всех подстановок s из $S(A, Y)$ на число всех подстановок t из $S(B—A, B—Y)$. Число множеств $F(Y)$ равно числу $\binom{3}{2}$ множеств Y в классе $\mathscr{P}_2(\{1, 2, 3\})$. В соответствии с правилом сложения из сказанного следует, что

$$\binom{3}{2}2!\,(3-2)!=3!,\quad \binom{3}{2}=\frac{3!}{2!\,(3-2)!}=3.$$

Таким образом, выбрать две из трех различных монет можно $\binom{3}{2}=3$ способами.

Пример 2. *Сколькими способами можно выбрать двух из четырех гостей $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$, $\mathscr{D}$?*

Каждую выборку $m=2$ из $n=4$ гостей $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$, $\mathscr{D}$ можно описать частью Y множества $B=\{\mathscr{A}, \mathscr{B}, \mathscr{C}, \mathscr{D}\}$, составленной из $m=2$ букв, обозначающих выбранных гостей. Эти части составляют класс

$$\mathscr{P}_2(\{\mathscr{A}, \mathscr{B}, \mathscr{C}, \mathscr{D}\})=\{\{\mathscr{A}, \mathscr{B}\}, \{\mathscr{A}, \mathscr{C}\}, \{\mathscr{A}, \mathscr{D}\},$$

$$\{\mathscr{B}, \mathscr{C}\}, \{\mathscr{B}, \mathscr{D}\}, \{\mathscr{C}, \mathscr{D}\}\}.$$

Как и в примере 1, существует зависимость между числом $\binom{4}{2}$ множеств этого класса и числом 4! перестановок множества $B=\{\mathscr{A}, \mathscr{B}, \mathscr{C}, \mathscr{D}\}$. Каждые подстановка s вместо $A=\{\mathscr{A}, \mathscr{B}\}$ множества Y и подстановка t вместо $B-A=\{\mathscr{C}, \mathscr{D}\}$ множества $B-Y$ определяют перестановку u множества $B=\{\mathscr{A}, \mathscr{B}, \mathscr{C}, \mathscr{D}\}$:

Y	s	t	u
$\{\mathscr{A}, \mathscr{B}\}$	$\begin{pmatrix}\mathscr{A} & \mathscr{B}\\ \mathscr{A} & \mathscr{B}\end{pmatrix}$	$\begin{pmatrix}\mathscr{C} & \mathscr{D}\\ \mathscr{C} & \mathscr{D}\end{pmatrix}$	$\begin{pmatrix}\mathscr{A} & \mathscr{B} & \mathscr{C} & \mathscr{D}\\ \mathscr{A} & \mathscr{B} & \mathscr{C} & \mathscr{D}\end{pmatrix}$
	$\begin{pmatrix}\mathscr{A} & \mathscr{B}\\ \mathscr{A} & \mathscr{B}\end{pmatrix}$	$\begin{pmatrix}\mathscr{C} & \mathscr{D}\\ \mathscr{D} & \mathscr{C}\end{pmatrix}$	$\begin{pmatrix}\mathscr{A} & \mathscr{B} & \mathscr{C} & \mathscr{D}\\ \mathscr{A} & \mathscr{B} & \mathscr{D} & \mathscr{C}\end{pmatrix}$
	$\begin{pmatrix}\mathscr{A} & \mathscr{B}\\ \mathscr{B} & \mathscr{A}\end{pmatrix}$	$\begin{pmatrix}\mathscr{C} & \mathscr{D}\\ \mathscr{C} & \mathscr{D}\end{pmatrix}$	$\begin{pmatrix}\mathscr{A} & \mathscr{B} & \mathscr{C} & \mathscr{D}\\ \mathscr{B} & \mathscr{A} & \mathscr{C} & \mathscr{D}\end{pmatrix}$
	$\begin{pmatrix}\mathscr{A} & \mathscr{B}\\ \mathscr{B} & \mathscr{A}\end{pmatrix}$	$\begin{pmatrix}\mathscr{C} & \mathscr{D}\\ \mathscr{D} & \mathscr{C}\end{pmatrix}$	$\begin{pmatrix}\mathscr{A} & \mathscr{B} & \mathscr{C} & \mathscr{D}\\ \mathscr{B} & \mathscr{A} & \mathscr{D} & \mathscr{C}\end{pmatrix}$
$\{\mathscr{A}, \mathscr{C}\}$	$\begin{pmatrix}\mathscr{A} & \mathscr{B}\\ \mathscr{A} & \mathscr{C}\end{pmatrix}$	$\begin{pmatrix}\mathscr{C} & \mathscr{D}\\ \mathscr{B} & \mathscr{D}\end{pmatrix}$	$\begin{pmatrix}\mathscr{A} & \mathscr{B} & \mathscr{C} & \mathscr{D}\\ \mathscr{A} & \mathscr{C} & \mathscr{B} & \mathscr{D}\end{pmatrix}$
	$\begin{pmatrix}\mathscr{A} & \mathscr{B}\\ \mathscr{A} & \mathscr{C}\end{pmatrix}$	$\begin{pmatrix}\mathscr{C} & \mathscr{D}\\ \mathscr{D} & \mathscr{B}\end{pmatrix}$	$\begin{pmatrix}\mathscr{A} & \mathscr{B} & \mathscr{C} & \mathscr{D}\\ \mathscr{A} & \mathscr{C} & \mathscr{D} & \mathscr{B}\end{pmatrix}$
	$\begin{pmatrix}\mathscr{A} & \mathscr{B}\\ \mathscr{C} & \mathscr{A}\end{pmatrix}$	$\begin{pmatrix}\mathscr{C} & \mathscr{D}\\ \mathscr{B} & \mathscr{D}\end{pmatrix}$	$\begin{pmatrix}\mathscr{A} & \mathscr{B} & \mathscr{C} & \mathscr{D}\\ \mathscr{C} & \mathscr{A} & \mathscr{B} & \mathscr{D}\end{pmatrix}$
	$\begin{pmatrix}\mathscr{A} & \mathscr{B}\\ \mathscr{C} & \mathscr{A}\end{pmatrix}$	$\begin{pmatrix}\mathscr{C} & \mathscr{D}\\ \mathscr{D} & \mathscr{B}\end{pmatrix}$	$\begin{pmatrix}\mathscr{A} & \mathscr{B} & \mathscr{C} & \mathscr{D}\\ \mathscr{C} & \mathscr{A} & \mathscr{D} & \mathscr{B}\end{pmatrix}$

Упражнение. Закончить составление таблицы.

Снова множества $F(Y)=S(A, Y)\times S(B-A, B-Y)$ пар подстановок s, t, описываемые средними клетками таблицы, не пересекаются и в сумме образуют множество, число элементов которого равно числу всех перестановок множества B. В соответствии с правилом умножения число пар подстановок s, t в каждом из множеств $F(Y)$ равно произведению $2!(4-2)!$ числа всех подстановок s из $S(A, Y)$ на число всех подстановок t из $S(B-A,$

$B - Y$). Число множеств $F(Y)$ равно числу $\binom{4}{2}$ множеств Y в классе $\mathcal{P}_2(\{\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}\})$. В соответствии с правилом сложения из сказанного следует, что

$$\binom{4}{2} 2!\,(4-2)! = 4!, \quad \binom{4}{2} = \frac{4!}{2!\,(4-2)!} = 6.$$

Таким образом, выбрать двух из четырех различных гостей можно $\binom{4}{2} = 6$ способами.

Пример 3. *Сколькими способами можно выбрать три из пяти букв слова camel?*

Каждую выборку $m=3$ из $n=5$ букв c, a, m, e, l можно описать частью Y множества $B=\{c, a, m, e, l\}$, составленной из выбирающихся $m=3$ букв. Эти части образуют класс

$$\mathcal{P}_3(\{c, a, m, e, l\}) = \{\{c, a, m\}, \{c, a, e\}, \{c, m, e\}, \{a, m, e\}, \{c, a, l\}, \{c, m, l\}, \{a, m, l\}, \{c, e, l\}, \{a, c, l\}, \{m, e, l\}\}.$$

Как и в примерах 1, 2, существует зависимость между числом $\binom{5}{3}$ множеств этого класса и числом 5! перестановок множества $B=\{c, a, m, e, l\}$. Каждые подстановка s вместо $A=\{c, a, m\}$ множества Y и подстановка t вместо $B-A$ множества $B-Y$ определяют перестановку u множества $B=\{c, a, m, e, l\}$:

Y	s	t	u
	$\begin{pmatrix} c & a & m \\ c & a & m \end{pmatrix}$	$\begin{pmatrix} e & l \\ e & l \end{pmatrix}$	$\begin{pmatrix} c & a & m & e & l \\ c & a & m & e & l \end{pmatrix}$
	$\begin{pmatrix} c & a & m \\ c & a & m \end{pmatrix}$	$\begin{pmatrix} e & l \\ l & e \end{pmatrix}$	$\begin{pmatrix} c & a & m & e & l \\ c & a & m & l & e \end{pmatrix}$
	$\begin{pmatrix} c & a & m \\ a & c & m \end{pmatrix}$	$\begin{pmatrix} e & l \\ e & l \end{pmatrix}$	$\begin{pmatrix} c & a & m & e & l \\ a & c & m & e & l \end{pmatrix}$
	$\begin{pmatrix} c & a & m \\ a & c & m \end{pmatrix}$	$\begin{pmatrix} e & l \\ l & e \end{pmatrix}$	$\begin{pmatrix} c & a & m & e & l \\ a & c & m & l & e \end{pmatrix}$
	$\begin{pmatrix} c & a & m \\ m & a & c \end{pmatrix}$	$\begin{pmatrix} e & l \\ e & l \end{pmatrix}$	$\begin{pmatrix} c & a & m & e & l \\ m & a & c & e & l \end{pmatrix}$
$\{c, a, m\}$	$\begin{pmatrix} c & a & m \\ m & a & c \end{pmatrix}$	$\begin{pmatrix} e & l \\ l & e \end{pmatrix}$	$\begin{pmatrix} c & a & m & e & l \\ m & a & c & l & e \end{pmatrix}$
	$\begin{pmatrix} c & a & m \\ c & m & a \end{pmatrix}$	$\begin{pmatrix} e & l \\ e & l \end{pmatrix}$	$\begin{pmatrix} c & a & m & e & l \\ c & m & a & e & l \end{pmatrix}$
	$\begin{pmatrix} c & a & m \\ c & m & a \end{pmatrix}$	$\begin{pmatrix} e & l \\ l & e \end{pmatrix}$	$\begin{pmatrix} c & a & m & e & l \\ c & m & a & l & e \end{pmatrix}$
	$\begin{pmatrix} c & a & m \\ m & c & a \end{pmatrix}$	$\begin{pmatrix} e & l \\ e & l \end{pmatrix}$	$\begin{pmatrix} c & a & m & e & l \\ m & c & a & e & l \end{pmatrix}$

Y	s	t	u
	$\begin{pmatrix} c & a & m \\ m & c & a \end{pmatrix}$	$\begin{pmatrix} e & l \\ l & e \end{pmatrix}$	$\begin{pmatrix} c & a & m & e & l \\ m & c & a & l & e \end{pmatrix}$
	$\begin{pmatrix} c & a & m \\ m & c & a \end{pmatrix}$	$\begin{pmatrix} e & l \\ e & l \end{pmatrix}$	$\begin{pmatrix} c & a & m & e & l \\ a & m & c & e & l \end{pmatrix}$
	$\begin{pmatrix} c & a & m \\ a & m & c \end{pmatrix}$	$\begin{pmatrix} e & l \\ l & e \end{pmatrix}$	$\begin{pmatrix} c & a & m & e & l \\ a & m & c & l & e \end{pmatrix}$
	$\begin{pmatrix} c & a & m \\ c & a & e \end{pmatrix}$	$\begin{pmatrix} e & l \\ m & l \end{pmatrix}$	$\begin{pmatrix} c & a & m & e & l \\ c & a & e & m & l \end{pmatrix}$
$\{c, a, e\}$	$\begin{pmatrix} c & a & m \\ c & a & e \end{pmatrix}$	$\begin{pmatrix} e & l \\ l & m \end{pmatrix}$	$\begin{pmatrix} c & a & m & e & l \\ c & a & e & l & m \end{pmatrix}$
	. . .	. . .	. . .

Множества $F(Y)=C(A,\ Y)\times S(B-A,\ B-Y)$ пар подстановок s, t, описываемые средними клетками таблицы, не пересекаются и в сумме образуют множество, число элементов которого равно числу 5! всех перестановок множества B. Число пар подстановок s, t в каждом из множеств $F(Y)$ равно произведению $3!(5-3)!$, а число множеств $F(Y)$ — числу $\binom{5}{3}$ множеств Y в классе $\mathscr{P}_3(\{c, a, m, e, l\})$. Поэтому

$$\binom{5}{3}3!\,(5-3)! = 5!,\quad \binom{5}{3} = \frac{5!}{3!\,(5-3)!} = 10.$$

Таким образом, выбрать три из пяти различных букв можно $\binom{5}{3} = 10$ способами.

2.3. Формула для числа выборок

Рассмотрим произвольное натуральное число n, натуральное число $m\leqslant n$ и множество B из n элементов. Каждую часть Y множества B, составленную из m элементов, условимся называть *выборкой* m из n элементов множества B. Класс всех таких выборок обозначим символом $\mathscr{P}_m(B)$. Число этих выборок, т. е. число множеств класса $\mathscr{P}_m(B)$, определяется числами m, n и обозначается символом $\binom{n}{m}$. Вместо *выборка* говорят также *сочетание.*

В частности, если $m=0$, класс $\mathscr{P}_0(B)$ состоит из пустого множества и $\binom{n}{0}=1$. Если $m=1$, то класс $\mathscr{P}_1(B)$ состоит из элементарных частей множества B и $\binom{n}{1}=n$.

Примеры 1, 2, 3 показывают, что

$$\binom{3}{2} = \frac{3!}{2!\,(3-2)!}, \quad \binom{4}{2} = \frac{4!}{2!\,(4-2)!},$$

$$\binom{5}{3} = \frac{5!}{3!\,(5-3)!}.$$

Эти равенства получены вследствие того, что каждая перестановка множества B эквивалентна выбору его части Y и подстановке вместо некоторой части A множества B этой части Y, а вместо части $B—A$ — части $B—Y$. Поэтому число $n!$ всех перестановок множества B равно произведению числа $\binom{n}{m}$ всех рассматриваемых выборок и чисел $m!, (n—m)!$ всех подстановок вместо A множества Y, а вместо $B—A$ — множества $B—Y$. Это равенство позволяет выразить число выборок через число перестановок. По такой схеме и доказывается общая

Формула для числа выборок:

$$\binom{n}{m} = \frac{n!}{m!\,(n-m)!}.$$

— Рассмотрим произвольное множество A класса $\mathcal{P}_m(B)$. Каждое множество Y этого класса определяет множество

$$F(Y) = S(A,\ Y) \times S(B—A,\ B—Y)$$

пар s, t подстановок s вместо A множества Y и подстановок t вместо $B—A$ множества $B—Y$.

Так как множества $S(A,\ Y)$ попарно не пересекаются, то и множества $F(Y)$ попарно не пересекаются.

Каждая пара подстановок s, t из каждого множества $F(Y)$ определяет перестановку u множества B со значением $u(x) = s(x)$ для каждого $x \in A$ и значением $u(x) = t(x)$ для каждого $x \in B—A$. Различные пары подстановок s, t определяют по этому правилу различные перестановки u. И каждая перестановка u множества B определяется парой s, t из множества $F[u(A)]$, образованной сужением s на A и t на $B—A$ перестановки u. Следовательно, данное правило определяет взаимно однозначное отображение суммы множеств $F(Y)$ рассматриваемого семейства на множество $S(B,\ B)$ всех перестановок множества B, и эти множества имеют одно и то же число элементов $n!$.

По правилу умножения и по формуле для числа подстановок для каждого множества Y класса $\mathcal{P}_m(B)$ число пар подстановок s, t в множестве $F(Y)$ равно произведению $m!\,(n—m)!$. Из правила сложения и сказанного следует, что

$$n! = n\left(\sum F(Y)\right) = \sum n(F(Y)) = \binom{n}{m} m!\,(n-m)!.$$

Значит,

$$\binom{n}{m} = \frac{n!}{m!\,(n-m)!}.$$

Формула для числа выборок доказана. +

В частности, для каждого натурального числа n верны равенства

$$\binom{n}{0} = \frac{n!}{0!\,(n-0)!} = \frac{n!}{n!} = 1,$$

$$\binom{n}{1} = \frac{n!}{1!\,(n-1)!} = \frac{n!}{(n-1)!} = n,$$

$$\binom{n}{2} = \frac{n!}{2!\,(n-2)!} = \frac{n(n-1)}{2}.$$

Формула для числа выборок дает решение классической задачи о числе выборок: *m из n различных предметов можно выбрать $\binom{n}{m} = n!/m!\,(n-m)!$ способами.*

§ 4. РАЗМЕЩЕНИЯ

Классической задачей комбинаторики является задача о числе размещений, содержание которой можно выразить вопросом: ***скольким и способами можно выбрать и разместить по m различным местам m из n различных предметов***?

1.4. Примеры

Рассмотрим несколько примеров.

Пример 1. *Сколькими способами можно выбрать и разместить по двум местам* 1, 2 *две из трех монет 1, 2, 3 копейки*?

Каждые такие выбор и размещение можно описать выборкой Y двух из трех элементов множества $B = \{1, 2, 3\}$ и подстановкой s вместо множества $A = \{1, 2\}$ выборки Y: на место $x \in A$ размещается монета $s(x) \in Y$. Выборки Y и подстановки s составляют первый и второй столбцы примера 1 из § 3. Рассматриваемые $\binom{3}{2} = 3$ выборки Y образуют класс $\mathscr{P}_2(B)$. Каждая выборка Y определяет множество $S(A, Y)$ из $2! = 2$ подстановок s. Всего в этих множествах

$$3^{(2)} = \binom{3}{2} 2! = \frac{3!}{(3-2)!} = 3 \cdot 2 = 6$$

подстановок.

Таким образом, выбрать и разместить по 2 различным местам 2 из 3 различных монет можно 6 способами.

Пример 2. *Сколькими способами можно выбрать и разместить по двум местам 1, 2 двух из четырех гостей* $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$, $\mathscr{D}$?

Каждые такие выбор и размещение можно описать выборкой Y двух из четырех элементов множества $B=\{\mathscr{A}, \mathscr{B}, \mathscr{C}, \mathscr{D}\}$ и подстановкой s вместо множества $A=\{1, 2\}$ выборки Y, на место $x \in A$ размещается гость $s(x) \in Y$.

Рассматриваемые $\binom{4}{2}=6$ выборок Y образуют класс

$$\mathscr{P}_2(B)=\{\{\mathscr{A}, \mathscr{B}\}, \{\mathscr{A}, \mathscr{C}\}, \{\mathscr{A}, \mathscr{D}\}, \{\mathscr{B}, \mathscr{C}\}, \{\mathscr{B}, \mathscr{D}\}, \{\mathscr{C}, \mathscr{D}\}\}.$$

Каждая выборка Y определяет множество $S(A, Y)$ из $2!=2$ подстановок s:

$$S(A, \{\mathscr{A}, \mathscr{B}\})=\left\{\begin{pmatrix} 1 & 2 \\ \mathscr{A} & \mathscr{B} \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ \mathscr{B} & \mathscr{A} \end{pmatrix}\right\},$$

$$S(A, \{\mathscr{A}, \mathscr{C}\})=\left\{\begin{pmatrix} 1 & 2 \\ \mathscr{A} & \mathscr{C} \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ \mathscr{C} & \mathscr{A} \end{pmatrix}\right\}, \ldots$$

Всего в этих множествах

$$4^{(2)}=\binom{4}{2}2!=\frac{4!}{(4-2)!}=4 \cdot 3=12$$

подстановок.

Таким образом, выбрать и разместить по 2 различным местам 2 из 4 гостей можно 12 способами.

Пример 3. *Сколько трехбуквенных слов можно составить из букв слова camel?*

Составление каждого такого слова сводится к выбору и размещению по трем местам 1, 2, 3 трех из пяти букв c, a, m, e, l. Эти выбор и размещение можно описать выборкой Y трех из пяти элементов множества $B=\{c, a, m, e, l\}$ и подстановкой s вместо множества $A=\{1, 2, 3\}$ выборки Y: на место $x \in A$ ставится буква $s(x) \in Y$.

Рассматриваемые $\binom{5}{3}=10$ выборок Y образуют класс $\mathscr{P}_3(B)$. Каждая выборка Y определяет множество $S(A, Y)$ из $3!=6$ подстановок s. Всего в этих множествах

$$5^{(3)}=\binom{5}{3}3!=\frac{5!}{(5-3)!}=5 \cdot 4 \cdot 3=60$$

подстановок.

Таким образом, из букв слова *camel* можно составить 60 трехбуквенных слов.

2.4. Формула для числа размещений

Рассмотрим произвольные натуральные числа n и $m \leqslant n$, множество A из m элементов и множество B из n элементов, произвольную часть Y множества B, составленную из m элемен-

тов. Каждую подстановку s вместо A некоторой части Y множества B назовем *размещением* множества A по B. Число всех таких размещений, т. е. число элементов суммы $\Sigma S(A, Y)$ всех множеств $S(A, Y)$, определяется числами n, m и обозначается символом $n^{(m)}$.

В частности, если $m=0$, то множество размещений состоит из пустой подстановки и

$$n^{(0)}=1.$$

Если $m=1$, множество размещений состоит из подстановок элементарных частей множества B и

$$n^{(1)}=n.$$

Примеры 1—3 показывают, что

$$3^{(2)}=\frac{3!}{(3-2)!},\quad 4^{(2)}=\frac{4!}{(4-2)!},\quad 5^{(3)}=\frac{5!}{(5-3)!}.$$

Эти равенства получены непосредственно из определений размещения через выборку и подстановку по формуле для числа выборок. Используя эту формулу, легко убеждаемся в том, что число $n^{(m)}$ размещений множества A из m элементов по множеству B из n элементов выражает

Формула для числа размещений:

$$n^{(m)}=\frac{n!}{(n-m)!}.$$

— По данному определению размещения задача сводится к вычислению числа элементов суммы $\Sigma S(A, Y)$ множеств $S(A, Y)$ подстановок вместо множества A составленной из m элементов части Y множества B. Так как множества $S(A, Y)$ попарно не пересекаются, то по правилу сложения

$$n^{(m)}=n(\Sigma S(A, Y))=\Sigma n(S(A, Y)).$$

В последней сумме каждое из слагаемых равно $m!$, а число слагаемых — $\binom{n}{m}$. По формуле для числа выборок отсюда следует, что

$$n^{(m)}=\binom{n}{m}m!=\frac{n!m!}{m!(n-m)!}=\frac{n!}{(n-m)!}.$$

Формула для числа размещений доказана. +

Из формул для числа размещений и числа подстановок следует, что при $m\geqslant 2$ число размещений равно произведению чисел $n, \ldots, n-m+1$:

$$n^{(m)}=n\cdot\ \ldots\ \cdot(n-m+1).$$

В частности,

$$n^{(2)}=n(n-1),\quad n^{(3)}=n(n-1)(n-2).$$

Формула для числа размещений дает решение классической задачи о числе размещений: *выбрать и разместить по различным местам* m *из* n *различных предметов можно* $n^{(m)}=n!/(n-m)!$ *способами.*

При $m=n$ задача о числе размещений сводится к задаче о числе перестановок. В соответствии с этим верно равенство $n^{(n)}=n!$.

Задачи к § 4 собраны в § 4 главы 1 части II.

§ 5. ФОРМУЛА НЬЮТОНА ДЛЯ БИНОМА

Знаменитая биномиальная формула Ньютона используется в самых различных рассуждениях. Ее естественным обобщением является часто применяемая полиномиальная формула.

1.5. Правило симметрии

В биномиальной формуле используются коэффициенты $\binom{n}{m}$. При действиях с ними бывает полезно

Правило симметрии:

$$\binom{n}{m}=\binom{n}{n-m}.$$

Это правило непосредственно следует из формулы для числа выборок. Оно выражает тот факт, что каждая выборка Y каких-нибудь m из n элементов множества B определяет свое дополнение $B-Y$, состоящее из $n-m$ элементов.

Описательно правило симметрии означает следующее: безразлично, какие из m и $n-m$ рассматриваемых предметов считать выбранными, а какие — оставшимися. В частности, в примерах 1 и 3 § 3 было бы удобнее считать выбранными соответственно одну монету и две буквы, игравшие роль оставшихся.

2.5. Правило Паскаля

При подсчете биномиальных коэффициентов $\binom{n}{m}$ иногда удобно использовать правило, сформулированное Паскалем.

В соответствии с определением выборки как части множества можно употреблять символ для числа выборок при каждом целом m:

$$\binom{n}{m}=\begin{cases} 0 & (m<0), \\ \dfrac{n!}{m!\,(n-m)!} & (0\leqslant m\leqslant n), \\ 0 & (n<m). \end{cases}$$

Равенства нулю при $m<0$ и $n<m$ соответствуют тому, что множество из n элементов не содержит частей со строго отрицательным числом элементов или с числом элементов, строго бóльшим n.

Правило Паскаля:

$$\binom{n+1}{m}=\binom{n}{m}+\binom{n}{m-1}.$$

— Если $m<0$, то обе части равенства Паскаля равны нулю. Если $m=0$, то оно сводится к равенству $1=1+0$, если $m=n+1$, — к равенству $1=0+1$. Если $m>n+1$, то обе части равенства Паскаля равны нулю.

При $0<m<n+1$ с помощью формулы для числа выборок получаем

$$\binom{n}{m}+\binom{n}{m-1}=\frac{n!}{m!\,(n-m)!}+\frac{n!}{(m-1)!\,(n-m+1)!}=$$

$$=\frac{n!}{m!\,(n-m+1)!}(n-m+1+m)=\frac{n!\,(n+1)}{m!\,(n-m+1)!}=$$

$$=\frac{(n+1)!}{m!\,(n+1-m)!}=\binom{n+1}{m}.$$

Правило Паскаля верно и в этом случае. +

При $0<m\leqslant n$ правило Паскаля выражает тот факт, что каждая выборка m элементов из $n+1$ либо содержит некоторый данный элемент, либо нет, причем число выборок первого типа равно числу выборок m—1 элементов из n остальных, а число выборок второго типа — числу выборок m элементов из n.

Правило Паскаля наглядно поясняется треугольником Паскаля:

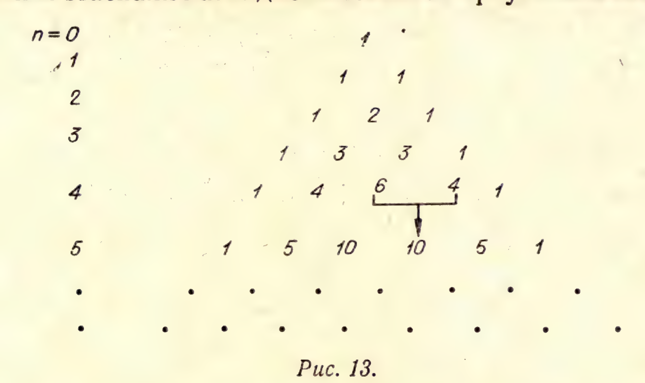


Рис. 13.

На m-м месте n-й строки этого треугольника расположено число $\binom{n}{m}$. Начальная строка треугольника и начальное место каждой строки считаются нулевыми. Боковые стороны треугольника Паскаля образованы числами 1. А каждое m-е внутреннее

число $(n+1)$-й строки $(0<m<n+1)$ в соответствии с правилом Паскаля равно сумме расположенных над ним $(m-1)$-го и m-го чисел n-й строки. Например, отмеченное на рисунке равенство $10=4+6$ соответствует равенству $\binom{5}{3}=\binom{4}{3}+\binom{4}{2}$.

Упражнение. Выписать 10-ю строку треугольника Паскаля.

3.5. Предварительное обсуждение формулы Ньютона

Заметим, что в $n=0$, 1, 2, 3-й строке треугольника Паскаля $m=0,\ldots, n$-е число $\binom{n}{m}$ является коэффициентом при $x^m y^{n-m}$ в равенстве для n-й степени $(x+y)^n$ суммы $x+y$ произвольных чисел x и y:

$$(x+y)^0=\binom{0}{0}x^0y^{0-0}$$
$$(x+y)^1=\binom{1}{0}x^0y^{1-0}+\binom{1}{1}x^1y^{1-1}$$
$$(x+y)^2=\binom{2}{0}x^0y^{2-0}+\binom{2}{1}x^1y^{2-1}+\binom{2}{2}x^2y^{2-2}$$
$$(x+y)^3=\binom{3}{0}x^0y^{3-0}+\binom{3}{1}x^1y^{3-1}+\binom{3}{2}x^2y^{3-2}+\binom{3}{3}x^3y^{3-3}$$

Эта закономерность имеет общий характер и поясняется следующим рассуждением.

Ясно, что $(x+y)^n$ является полиномом n-й степени относительно переменной x:

$$(x+y)^n=(x+y)\ldots(x+y)=a_0x^0+a_1x^1+\ldots+a_mx^m+\ldots+a_nx^n.$$

Раскрывая скобки в произведении n сумм $x+y$, можно получить m-ю степень x^m, если, и только если в каких-либо m суммах $x+y$ выбрать в качестве множителей слагаемое x, а в остальных $n-m$ суммах — слагаемое y. Такой выбор осуществляется $\binom{n}{m}$ способами. Следовательно,

$$a_m=\binom{n}{m}y^{n-m}$$

и

$$(x+y)^n=\binom{n}{0}x^0y^{n-0}+\binom{n}{1}x^1y^{n-1}+\ldots+\binom{n}{m}x^my^{n-m}+\ldots+$$
$$+\binom{n}{n}x^ny^{n-n}=y^n+nxy^{n-1}+\ldots+\frac{n!}{m!\,(n-m)!}x^my^{n-m}+\ldots+x^n.$$

Коэффициенты при $x^m y^{n-m}$ в правой части составляют строку треугольника Паскаля. В частности,

$$(x+y)^4 = \binom{4}{0} x^0 y^{4-0} + \binom{4}{1} x^1 y^{4-1} + \binom{4}{2} x^2 y^{4-2} + \binom{4}{3} x^3 y^{4-3} +$$

$$+ \binom{4}{4} x^4 y^{4-4} = 1y^4 + 4xy^3 + 6x^2y^2 + 4x^3y + 1x^4,$$

$$(x+y)^5 = \binom{5}{0} x^0 y^{5-0} + \binom{5}{1} x^1 y^{5-1} + \binom{5}{2} x^2 y^{5-2} + \binom{5}{3} x^3 y^{5-3} +$$

$$+ \binom{5}{4} x^4 y^{5-4} + \binom{5}{5} x^5 y^{5-5} = 1y^5 + 5xy^4 + 10x^2y^3 +$$

$$+ 10x^3y^2 + 5x^4y + 1x^5.$$

Найденные для $(x+y)^n$ равенства являются развернутыми выражениями биномиальной формулы.

4.5. Доказательство формулы Ньютона

Приведем сокращенную запись и доказательство биномиальной формулы, основанное на правиле Паскаля и принципе индукции. Доказательство повторяет рассуждения пункта 3.5.

Рассмотрим произвольные числа x, y и натуральное число n. Условимся считать, что нулевая степень каждого числа равна единице:

$$z^0 = 1$$

для каждого числа z (в частности, для числа 0).

Формула Ньютона:

$$(x+y)^n = \sum_{0 \leqslant m \leqslant n} \binom{n}{m} x^m y^{n-m}.$$

— Рассмотрим множество M всех натуральных чисел n, для которых такое равенство верно при каждых числах x и y.

1. Число 0 принадлежит множеству M, так как при $n=0$ обе части рассматриваемого равенства обращаются в 1.

2. Для каждого натурального числа n верно предложение: если $n \in M$, то $(n+1) \in M$.

Действительно, если $n \in M$, то

$$(x+y)^{n+1} = (x+y)(x+y)^n = (x+y) \sum_{0 \leqslant m \leqslant n} \binom{n}{m} x^m y^{n-m} =$$

$$= \sum_{0 \leqslant m \leqslant n} \binom{n}{m} x^{m+1} y^{n-m} + \sum_{0 \leqslant m \leqslant n} \binom{n}{m} x^m y^{(n+1)-m}.$$

Заметим, что

$$\binom{n}{-1} = 0, \quad \binom{n}{n+1} = 0.$$

Добавим к первой из полученных сумм равное нулю слагаемое, соответствующее значению —1 индекса суммирования. Затем изменим интервал суммирования, перейдя снова к положительным значениям индекса. Рассматриваемая сумма при этом не изменится:

$$\sum_{0\leqslant m\leqslant n} \binom{n}{m} x^{m+1}y^{n-m} \quad \sum_{-1\leqslant l\leqslant n} \binom{n}{l} x^{l+1}y^{n-l} =$$

$$= \sum_{0\leqslant l+1\leqslant n+1} \binom{n}{(l+1)-1} x^{l+1}y^{(n+1)-(l+1)} =$$

$$= \sum_{0\leqslant m\leqslant n+1} \binom{n}{m-1} x^{m}y^{(n+1)-m}.$$

Аналогично

$$\sum_{0\leqslant m\leqslant n} \binom{n}{m} x^{m}y^{(n+1)-m} = \sum_{0\leqslant m\leqslant n+1} \binom{n}{m} x^{m}y^{(n+1)-m}.$$

Используя эти равенства и правило Паскаля, получаем:

$$(x+y)^{n+1} = \sum_{0\leqslant m\leqslant n+1} \binom{n}{m-1} x^{m}y^{(n+1)-m} + \sum_{0\leqslant m\leqslant n+1} \binom{n}{m} x^{m}y^{(n+1)-m} =$$

$$= \sum_{0\leqslant m\leqslant n+1} \left[\binom{n}{m-1} + \binom{n}{m}\right] x^{m}y^{(n+1)-m} =$$

$$= \sum_{0\leqslant m\leqslant n+1} \binom{n+1}{m} x^{m}y^{(m+1)-m}.$$

Значит, $n+1 \in M$.

По принципу индукции из пунктов 1 и 2 следует, что множество M совпадает с множеством N всех натуральных чисел, т. е. доказываемое равенство верно для каждого натурального числа n и произвольных чисел x, y. +

5.5. Неравенство Бернулли

Используя формулу Ньютона, легко доказать, что для каждых числа $c>1$ и натурального числа $n>1$ верно классическое **Неравенство Бернулли:**

$$c^n > 1 + n(c-1).$$

— По формуле Ньютона для каждого натурального числа n и чисел $x=c-1$, $y=1$ верны равенства

$$c^n = (x+1)^n = \sum_{0\leqslant m\leqslant n} \binom{n}{m} x^m.$$

По условию, $x>0$ и $n\geqslant 2$. Следовательно, каждое слагаемое в полученной сумме строго положительно и в ней по меньшей мере три слагаемых. Значит,

$$\sum_{0\leqslant m\leqslant n}\binom{n}{m}x^m \geqslant 1+nx+\frac{n(n-1)}{2}x^2>1+nx$$

и неравенство Бернулли верно. +

Пример. $\left(1+\frac{1}{100}\right)^{1000}>1+1000\frac{1}{100}>10.$

6.5*. Полиномиальная формула

Естественным обобщением формулы Ньютона является формула для степени суммы нескольких слагаемых.

Рассмотрим произвольные натуральное число n, натуральное число $k\geqslant 2$ и числа $x_1, \ldots, x_k$. Условимся символами $n_1, \ldots, n_k$ обозначать произвольные натуральные числа, в сумме равные n:

$$n_1+\ldots+n_k=n.$$

Символы $n_1, \ldots, n_k$ под знаком суммы означают, что складываются числа, соответствующие каждым натуральным числам $n_1, \ldots, n_k$, в сумме равным n.

Используя предлагаемую символику и формулу для числа выборок, формулу Ньютона можно записать следующим образом:

$$(x_1+x_2)^n=\sum_{n_1,n_2}\frac{n!}{n_1!\,n_2!}x_1^{n_1}x_2^{n_2}.$$

Верна обобщающая ее

Полиномиальная формула:

$$(x_1+\ldots+x_k)^n=\sum_{n_1,\ldots,n_k}\frac{n!}{n_1!\ldots n_k!}x_1^{n_1}\ldots x_k^{n_k}.$$

— Рассмотрим множество M всех натуральных чисел $k-2$ таких, что для числа k доказываемое равенство верно при каждых числах $x_1, \ldots, x_k$ и натуральном числе n.

1. Число 0 принадлежит множеству M, так как при $k=2$ доказываемая формула эквивалентна формуле Ньютона.

2. Для каждого натурального числа l верно предложение: если $l\in M$, то $l+1\in M$.

Рассмотрим произвольное натуральное число l и число $k=l+2$. Если $l\in M$, то с помощью формулы Ньютона для каждых чисел $x_1, \ldots, x_k, x_{k+1}$ и натурального числа n получаем:

$$(x_1+\ldots+x_k+x_{k+1})^n=\sum_{0\leqslant m\leqslant n}\frac{n!}{m!\,(n-m)!}(x_1+\ldots+x_k)^m x_{k+1}^{n-m}=$$

$$= \sum_{0 \leqslant m \leqslant n} \sum_{m_1 \dots m_k} \frac{n!}{m!\,(n-m)!} \frac{m!}{m_1! \dots m_k!} x_1^{m_1} \dots x_k^{m_k} x_{k+1}^{n-m} =$$

$$= \sum_{n_1, \dots, n_k,\, n_{k+1}} \frac{n!}{n_1! \dots n_k!\, n_{k+1}!} x_1^{n_1} \dots x_k^{n_k} x_{k+1}^{n_{k+1}}.$$

Значит, $l+1 = (k+1)-2 \in M$.

По принципу индукции из пунктов 1 и 2 следует, что множество M совпадает с множеством N всех натуральных чисел, т. е. доказываемое равенство верно для произвольных натуральных чисел n, $k \geqslant 2$ и чисел $x_1, \dots, x_k$. +

Пример. Для каждых чисел x_1, x_2, x_3 верны равенства

$$(x_1 + x_2 + x_3)^2 = \frac{2!}{2!\,0!\,0!} x_1^2 x_2^0 x_3^0 + \frac{2!}{0!\,2!\,0!} x_1^0 x_2^2 x_3^0 + \frac{2!}{0!\,0!\,2!} x_1^0 x_2^0 x_3^2 +$$

$$+ \frac{2!}{1!\,1!\,0!} x_1^1 x_2^1 x_3^0 + \frac{2!}{1!\,0!\,1!} x_1^1 x_2^0 x_3^1 + \frac{2!}{0!\,1!\,1!} x_1^0 x_2^1 x_3^1 = x_1^2 + x_2^2 + x_3^2 +$$

$$+ 2x_1x_2 + 2x_1x_3 + 2x_2x_3.$$

§ 6*. РАЗБИЕНИЯ

Классическую задачу о числе выборок можно сформулировать эквивалентным образом: сколькими способами можно разбить n различных предметов на $k=2$ группы 1, 2 так, чтобы в первой было $n_1=m$, а во второй — $n_2=n-m$ предметов? Естественно сформулировать такую задачу для разбиений на несколько групп: *сколькими способами можно разбить n различных предметов на k групп $1, \dots, k$ по $n_1, \dots, n_k$ предметов?*

1.6. Примеры

Рассмотрим несколько примеров.

Пример 1. *Сколькими способами можно разбить три монеты 1, 2, 3 копейки на три группы 1, 2, 3 по одной монете?*

Каждое такое разбиение можно описать разбиением множества $B=\{1, 2, 3\}$ из $n=3$ элементов на $k+1=3$ попарно не пересекающиеся части B_1, B_2, B_3, состоящие каждая из $n_1=n_2=n_3=1$ элементов и в сумме равные B. При каждом таком множестве B_3 множества $A_1=B_1$, $A_2=B_2$ образуют разбиение из $m=2$ элементов множества $A=B-B_3$ на $k=2$ части, состоящие каждая из $m_1=$ $=n_1=1$, $m_2=n_2=1$ элементов. Число $\binom{2}{1,\,1}$ таких разбиений равно числу $\binom{2}{1}$ выборок 1 элемента из 2. Следовательно, число

$\binom{3}{1,\,1,\,1}$ всех разбиений трех монет на три нумерованные группы по одной монете выражается равенствами

$$\binom{3}{1,\,1,\,1} = \binom{3}{1}\binom{2}{1,\,1} = \binom{3}{1}\binom{2}{1} = \frac{3!}{1!\,2!}\cdot\frac{2!}{1!\,1!} = \frac{3!}{1!\,1!\,1!} = 6.$$

Таким образом, число рассматриваемых разбиений, как и следовало ожидать, равно числу перестановок трех монет по трем местам.

Пример 2. *Сколькими способами можно разбить четырех гостей $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$, $\mathscr{D}$ на две пары 1, 2?*

Каждое такое разбиение можно описать выборкой $m=2$ из $n=4$ элементов множества $B=\{\mathscr{A}, \mathscr{B}, \mathscr{C}, \mathscr{D}\}$. Поэтому число $\binom{4}{2,\,2}$ рассматриваемых разбиений равно числу $\binom{4}{2}$ таких выборок:

$$\binom{4}{2,\,2} = \binom{4}{2} = \frac{4!}{2!\,2!} = 6.$$

Таким образом, разбить четырех гостей на первую и вторую пары можно шестью способами:

1	$\{\mathscr{A}, \mathscr{B}\}$	$\{\mathscr{A}, \mathscr{C}\}$	$\{\mathscr{A}, \mathscr{D}\}$	$\{\mathscr{B}, \mathscr{C}\}$	$\{\mathscr{B}, \mathscr{D}\}$	$\{\mathscr{C}, \mathscr{D}\}$
2	$\{\mathscr{C}, \mathscr{D}\}$	$\{\mathscr{B}, \mathscr{D}\}$	$\{\mathscr{B}, \mathscr{C}\}$	$\{\mathscr{A}, \mathscr{D}\}$	$\{\mathscr{A}, \mathscr{C}\}$	$\{\mathscr{A}, \mathscr{B}\}$

Таблица показывает, что нумерация пар существенна: симметричные относительно середины разбиения отличаются только номерами пар.

Пример 3. *Сколькими способами можно разбить пять букв слова camel на четыре группы 1, 2, 3, 4 так, чтобы в группах 1, 2, 3 было по одной букве, а в группе 4 — две?*

Каждое такое разбиение можно описать разбиением множества $B=\{c, a, m, e, l\}$ из $n=5$ элементов на $k+1=4$ попарно не пересекающиеся части B_1, B_2, B_3, B_4, состоящие соответственно из $n_1=n_2=n_3=1$, $n_4=2$ элементов и в сумме равные B. При каждом таком множестве B_4 множества $A_1=B_1$, $A_2=B_2$, $A_3=B_3$ образуют разбиение множества $A=B-B_4$ на $k=3$ части, состоящие каждая из $m_1=n_1=1$, $m_2=n_2=1$, $m_3=n_3=1$ элементов. Число $\binom{3}{1,\,1,\,1}$ таких разбиений равно подсчитанному в примере 1. Следовательно, число $\binom{5}{1,\,1,\,1,\,2}$ всех разбиений пяти букв на четыре группы

1, 2, 3, 4 соответственно из $n_1=1$, $n_2=1$, $n_3=1$, $n_4=2$ букв выражается равенствами

$$\binom{5}{1,\ 1,\ 1,\ 2}=\binom{5}{2}\binom{3}{1,\ 1,\ 1}=\frac{5!}{2!\,3!}\cdot\frac{3!}{1!\,1!\,1!}=\frac{5!}{1!\,1!\,1!\,2!}=60.$$

Таким образом, разбить пять букв на четыре нумерованные группы, первые три из которых состоят из одной, а четвертая — из двух букв, можно шестьюдесятью способами.

2.6. Формула для числа разбиений

Рассмотрим произвольные натуральные числа n, $k\geqslant 2$ и $n_1, \ldots, n_k$, в сумме равные n:

$$n_1+\ldots+n_k=n.$$

Рассмотрим произвольные множество A из n элементов и попарно не пересекающиеся множества $A_1, \ldots, A_k$ соответственно из $n_1, \ldots, n_k$ элементов, в сумме равные A:

$$A_1+\ldots+A_k=A.$$

Условимся называть каждое такое семейство $A_1, \ldots, A_k$ множеств $n_1, \ldots, n_k$*-разбиением* множества A. Число всех $n_1, \ldots, n_k$-разбиений определяется числами n, n_1, ..., n_k и обозначается символом $\binom{n}{n_1,\ \ldots,\ n_k}$.

В частности, если $k=2$, то каждое n_1, n_2-разбиение A_1, A_2 множества A из $n_1+n_2=n$ элементов определяется выборкой A_1 некоторых n_1 из n элементов множества A. По формуле для числа выборок отсюда следует, что

$$\binom{n}{n_1,\ n_2}=\frac{n!}{n_1!\,n_2!}.$$

Примеры 1—3 показывают, что

$$\binom{3}{1,\ 1,\ 1}=\frac{3!}{1!\,1!\,1!},\quad \binom{4}{2,\ 2}=\frac{4!}{2!\,2!},\quad \binom{5}{1,\ 1,\ 1,\ 2}=\frac{5!}{1!\,1!\,1!\,2!}.$$

Эти равенства были получены с помощью формулы для числа выборок и индукции. По такой схеме получается и общая

Формула для числа разбиений:

$$\binom{n}{n_1,\ \ldots,\ n_k}=\frac{n!}{n_1!\ldots n_k!}.$$

— Рассмотрим множество M всех натуральных чисел $k-2$ таких, что для числа k такое равенство верно для каждых рассматриваемых чисел $n_1, \ldots, n_k$, n и множества A.

1. Число 0 принадлежит множеству M, так как при $k=2$ доказываемая формула эквивалентна формуле для числа выборок.

2. Для каждого натурального числа l верно предложение: если $l \in M$, то $l+1 \in M$.

Рассмотрим произвольные натуральные числа $n_1, \ldots, n_k, n_{k+1}$, в сумме равные n, и множество B из n элементов. Выборка Z каждых n_{k+1} из n элементов множества B определяет множество $R(Z)$ всех $n_1, \ldots, n_k, n_{k+1}$-разбиений $B_1, \ldots, B_k, B_{k+1}$ множества B, для которых $B_{k+1} = Z$.

Множества $R(Z)$ попарно не пересекаются и в сумме образуют множество R всех $n_1, \ldots, n_k, n_{k+1}$-разбиений множества B:

$$R = \Sigma R(Z).$$

Число множеств $R(Z)$ равно числу $\binom{n}{n_{k+1}}$ выборок Z произвольных n_{k+1} из n элементов множества B.

Каждое принадлежащее $R(Z)$ разбиение $B_1, \ldots, B_k, B_{k+1}$ множества B определяет $n_1, \ldots, n_k$-разбиение $B_1, \ldots, B_k, B_{k+1}$ множества $A = B - Z$ из $n - n_{k+1}$ элементов. Различные разбиения из $R(Z)$ определяют различные такие разбиения множества A. И каждое $n_1, \ldots, n_k$-разбиение $A_1, \ldots, A_k$ множества $A = B - Z$ определяется принадлежащим $R(Z)$ разбиением $A_1, \ldots, A_k, Z$ множества B. Следовательно, число разбиений в $R(Z)$ равно числу $\binom{n-n_{k+1}}{n_1, \ldots, n_k}$ всех $n_1, \ldots, n_k$-разбиений множества A из $n - n_{k+1}$ элементов.

По правилу сложения для числа элементов, из сказанного следует, что число всех разбиений из R равно произведению числа множеств $R(Z)$ на число разбиений в каждом из них:

$$\binom{n}{n_1, \ldots, n_k, n_{k+1}} = \binom{n}{n_{k+1}} \binom{n-n_{k+1}}{n_1, \ldots, n_k}.$$

Рассмотрим произвольное натуральное число l и число $k = l+2$. Если $l \in M$, то, используя только что доказанное равенство и формулу для числа выборок, получаем

$$\binom{n}{n_1, \ldots, n_k, n_{k+1}} = \frac{n!}{n_{k+1}!\,(n-n_{k+1})!} \cdot \frac{(n-n_{k+1})}{n_1! \ldots n_k!} = \frac{n!}{n_1! \ldots n_k!\, n_{k+1}!}$$

Значит, $l+1 = (k+1) - 2 \in M$.

По принципу индукции из пунктов 1 и 2 следует, что множество M совпадает с множеством N всех натуральных чисел, т. е. доказываемое равенство верно для произвольных натуральных чисел $k \geqslant 2$ и $n_1 + \ldots + n_k = n$. +

Формула для числа выборок дает решение классической задачи о числе разбиений: *n различных предметов на k групп $1, \ldots, k$ по $n_1, \ldots, n_k$ предметов можно разбить $n!/n_1! \ldots n_k!$ способами.*

В частности, если в каждой группе ровно по одному предмету, то разбиения сводятся к перестановкам:

$$\binom{n}{1, \ldots, 1} = \frac{n!}{1! \ldots 1!} = n!.$$

Замечание. Формула для числа разбиений позволяет записать доказанную в предыдущем параграфе полиномиальную формулу в следующем эквивалентном виде:

$$(x_1 + \ldots + x_k)^n = \sum_{n_1, \ldots, n_k} \binom{n}{n_1, \ldots, n_k} x_1^{n_1} \ldots x_k^{n_k}.$$

Появление чисел $\binom{n}{n_1, \ldots, n_k}$ в качестве коэффициентов при $x_1^{n_1} \ldots x_k^{n_k}$ не случайно: раскрывая скобки в произведении n сумм $x_1 + \ldots + x_k$, получить $x_1^{n_1} \ldots x_k^{n_k}$ можно, если и только если в n_1 суммах $x_1 + \ldots + x_k$ выбрать в качестве множителей слагаемое x_1, ..., в n_k суммах — слагаемое x_k. Это можно сделать $\binom{n}{n_1, \ldots, n_k}$ способами.

§ 7*. ПЕРЕСТАНОВКИ С ПОВТОРЕНИЯМИ

В обсуждавшихся до сих пор задачах о числе перестановок, выборок, размещений и разбиений рассматривались различные предметы. Естественно сформулировать аналогичные задачи и для случая, когда среди предметов есть одинаковые. Тогда задачу о числе перестановок можно выразить вопросом: *сколькими способами можно переставить n предметов k различных типов $1, \ldots, k$ по $n_1, \ldots n_k$ одинаковых предметов, расположенных на n различных местах $1, \ldots, n$?*

1.7. Примеры

Рассмотрим несколько примеров.

Пример 1. *Сколькими способами можно переставить три монеты 1, 2, 3 копейки по трем различным местам?*

Рассматриваемые монеты различны. Такая задача была решена в § 2. Ответ: 3! способами.

Пример 2. *Сколькими способами можно пересадить близнецов из двух различных пар $\mathcal{A}$, $\mathcal{A}$ и $\mathcal{B}$, $\mathcal{B}$ на четырех местах 1, 2, 3, 4?*

Каждую пересадку близнецов $\mathcal{A}$, $\mathcal{A}$, $\mathcal{B}$, $\mathcal{B}$ на местах 1, 2, 3, 4 можно описать отображением f множества $A = \{1, 2, 3, 4\}$ в множество $B = \{\mathcal{A}, \mathcal{B}\}$: на место x пересаживается близнец $f(x)$. По

аналогии с подстановками такие отображения описываются таблицами

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ \mathscr{A} & \mathscr{A} & \mathscr{B} & \mathscr{B} \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ \mathscr{A} & \mathscr{B} & \mathscr{A} & \mathscr{B} \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ \mathscr{A} & \mathscr{B} & \mathscr{B} & \mathscr{A} \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ \mathscr{B} & \mathscr{A} & \mathscr{A} & \mathscr{B} \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ \mathscr{B} & \mathscr{A} & \mathscr{B} & \mathscr{A} \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ \mathscr{B} & \mathscr{B} & \mathscr{A} & \mathscr{A} \end{pmatrix}.$$

Каждое такое отображение f определяется 2, 2-разбиением $f^{-1}(\mathscr{A}), f^{-1}(\mathscr{B})$ множества A:

$$\{1, 2\}, \{3, 4\}; \{1, 3\}, \{2, 4\}; \{1, 4\}, \{2, 3\};$$
$$\{2, 3\}, \{1, 4\}; \{2, 4\}, \{1, 3\}; \{3, 4\}, \{1, 2\}.$$

Число отображений f равно числу

$$\binom{4}{2,\ 2} = \frac{4!}{2!\,2!} = 6$$

всех 2, 2-разбиений множества из 4 элементов.

Таким образом, каждая пересадка близнецов из двух пар на четырех местах эквивалентна разбиению множества этих четырех мест на два для первой и два для второй пары. Это разбиение можно провести шестью способами.

Пример 3. *Сколькими способами можно переставить буквы в слове canal?*

Каждую перестановку букв в слове *canal* можно описать некоторым отображением f множества $A = \{1, 2, 3, 4, 5\}$ в множество $B = \{c, n, l, a\}$: на место x ставится буква $f(x)$. Каждое такое отображение определяется 1, 1, 1, 2-разбиением $f^{-1}(c)$, $f^{-1}(n)$, $f^{-1}(l)$, $f^{-1}(a)$ множества A. Следовательно, число рассматриваемых отображений f равно числу

$$\binom{5}{1,\ 1,\ 1,\ 2} = \frac{5!}{1!\,1!\,1!\,2!} = 60$$

всех 1, 1, 1, 2-разбиений множества из 5 элементов.

Таким образом, каждая перестановка букв в слове *canal* эквивалентна разбиению множества пяти элементов на четыре группы, первые три из которых состоят из одного, а четвертая — из двух элементов. Это разбиение можно провести шестьюдесятью способами.

2.7. Формула для числа перестановок с повторениями

Рассмотрим произвольные натуральные числа $n \geqslant 1$, $k \geqslant 2$ и $n_1, \ldots, n_k$, в сумме равные n:

$$n_1 + \ldots + n_k = n.$$

Рассмотрим также множество

$$A = \{1, \ldots, n\}$$

и произвольное множество

$$B=\{b_1, \ldots, b_k\}$$

из k элементов, занумерованных в каком-либо порядке.

Каждое отображение f множества A в множество B, при котором прообразы $f^{-1}(b_1), \ldots, f^{-1}(b_k)$ элементов $b_1, \ldots, b_k$ состоят соответственно из $n_1, \ldots, n_k$ номеров:

$$n(f^{-1}(b_1))=n_1, \ldots, n(f^{-1}(b_k))=n_k,$$

условимся называть $n_1, \ldots, n_k$-*перестановкой* множества B.

В частности, если $n=k$ и $A=B$ при стандартной нумерации, то $1, \ldots 1$-перестановки являются обычными перестановками множества номеров $1, \ldots, n$.

Число всех $n_1, \ldots, n_k$-перестановок множества B, как и число всех $n_1, \ldots, n_k$-разбиений множества A, обозначим символом $\binom{n}{n_1, \ldots, n_k}$. Это оправдывает

Формула для числа перестановок с повторениями:

$$\binom{n}{n_1, \ldots, n_k}=\frac{n!}{n_1! \ldots n_k!}.$$

— По определению, каждая $n_1, \ldots, n_k$-перестановка множества B определяет $n_1, \ldots, n_k$-разбиение $f^{-1}(b_1), \ldots, f^{-1}(b_k)$ множества A. Различные перестановки f определяют различные разбиения множества A. И каждое $n_1, \ldots, n_k$-разбиение $A_1, \ldots, A_k$ множества A определяется $n_1, \ldots, n_k$-перестановкой f множества B со значениями $f(x)=b_1$ для каждого $x \in A_1, \ldots, f(x)=b_k$ для каждого $x \in A_k$. Следовательно, число $n_1, \ldots, n_k$-перестановок множества B равно числу $n_1, \ldots, n_k$-разбиений множества A, и доказываемая формула эквивалентна формуле для числа разбиений. +

Полученная формула для числа перестановок с повторениями позволяет решить сформулированную в начале параграфа задачу: *n предметов k различных типов $1, \ldots, k$ по $n_1, \ldots, n_k$ предметов, расположенных на n различных местах $1, \ldots, n$, можно переставить $n!/n_1! \ldots n_k!$ способами.*

3.7. Задача о числе отображений

Полиномиальная формула и формула для числа перестановок с повторениями позволяют ответить на вопрос: *сколько всего существует различных отображений множества $A=\{1, \ldots, n\}$ из n элементов в множество $B=\{1, \ldots, k\}$ из k элементов?*

При $x_1= \ldots =x_k=1$ из полиномиальной формулы вытекает

Следствие.

$$\sum_{n_1, \ldots, n_k}\binom{n}{n_1, \ldots, n_k}=k^n.$$

Каждое отображение f множества A в множество B является некоторой перестановкой множества B. Из формулы для числа перестановок и только что полученного равенства следует, что общее число отображений A в B равно k^n.

Таким образом, *существует всего k^n различных отображений множества A из n элементов в множество B из k элементов.*

§ 8*. ВЫБОРКИ С ПОВТОРЕНИЯМИ

Рассмотрим такую задачу о числе выборок с повторениями: *имеется по r одинаковых предметов каждого из n различных типов; сколькими способами можно выбрать m из этих $n \cdot r$ предметов?*

Если $r=1$, то предметы различны и задача сводится к уже решенной в § 3.

В этом параграфе рассматривается другой крайний случай, когда $r=m$. В этом случае выборка может даже полностью состоять из одинаковых предметов.

1.8. Примеры

Рассмотрим несколько примеров.

Пример 1. *Имеется по две монеты 1, 2, 3, копейки. Сколькими способами можно выбрать две из этих шести монет?*

Каждую такую выборку можно описать отображением f множества $A=\{1, 2, 3\}$ в множество $N=\{0, 1, 2, \ldots\}$ натуральных чисел: монет достоинством x копеек выбирается $f(x)$ штук. Используя стандартный порядок номеров 1, 2, 3, составляющих множество A, эти отображения можно описать строками, в которых на месте с номером x расположено соответствующее этому номеру значение $f(x)$ отображения f:

$$011, 002, 020, 101, 110, 200.$$

В частности, первая из этих строк описывает отображение f со значениями $f(1)=0$, $f(2)=1$, $f(3)=1$; вторая — со значениями

$$f(1)=0, \quad f(2)=0, \quad f(3)=2.$$

Будем натуральные числа 0, 1, 2, ... обозначать также соответственно символами 0, 01, 011, ... При этом условии рассматриваемые отображения f будут описываться также строками

$$00101, 00011, 00110, 01001, 01010, 01100.$$

Каждая из этих строк начинается с нуля и состоит из $n=3$ нулей и $m=2$ единиц. Поэтому она определяется выборкой $n-1=2$ номеров мест нулей из $n+m-1=3+2-1=4$ номеров 1, 2, 3, 4:

$$\{1, 3\}, \{1, 2\}, \{1, 4\}, \{2, 3\}, \{2, 4\} \{3, 4\}$$

(считается, что нуль в начале каждой строки занимает место с номером 0).

Таким образом, задача о числе выборок с повторениями $m=2$ из $n\cdot m=6$ монет $n=3$ видов сводится к задаче о числе выборок без повторений $n-1=3-1=2$ из $n+m-1=3+2-1=4$ номеров 1, 2, 3, 4. В соответствии с этим число интересующих нас выборок выражается равенствами

$$\binom{3+2-1}{3-1}=\binom{3+2-1}{2}=6.$$

Заметим, что в конечном счете каждому рассматриваемому отображению f была поставлена в соответствие выборка $Y=\{y_1, y_2\}$, элементы которой определяются равенствами

$$y_1=f(1)+1,\ y_2=f(1)+f(2)+2.$$

В частности, если $f(1)=0$, $f(2)=f(3)=1$, то $y_1=1$, $y_2=3$; если $f(1)=f(2)=0$, $f(3)=2$, то $y_1=1$, $y_2=2$.

Пример 2. *В гости приглашены четыре различные пары $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$, $\mathscr{D}$ близнецов. Сколькими способами можно выбрать двух из восьми гостей?*

Каждую такую выборку можно описать отображением f множества $A=\{\mathscr{A}, \mathscr{B}, \mathscr{C}, \mathscr{D}\}$ в множество $N=\{0, 1, 2, \ldots\}$ натуральных чисел: из пары x выбирается $f(x)$ близнецов. Используя алфавитный порядок букв, составляющих множество A, эти отображения можно описать строками, в которых на месте с порядковым номером буквы x в множестве A расположено соответствующее этой букве значение $f(x)$ отображения f:

$$0011,\quad 0002,\quad 0020,\quad 0101,\quad 0110,$$
$$0200,\quad 1001,\quad 1010,\quad 1100,\quad 2000.$$

В частности, первая из этих строк описывает отображение f со значениями $f(\mathscr{A})=0$, $f(\mathscr{B})=0$, $f(\mathscr{C})=1$, $f(\mathscr{D})=1$, вторая — со значениями $f(\mathscr{A})=0$, $f(\mathscr{B})=0$, $f(\mathscr{C})=0$, $f(\mathscr{D})=2$.

Будем натуральные числа 0, 1, 2, ... обозначать также соответственно символами 0, 01, 011,... При этом условии рассматриваемые отображения f будут описываться также строками

$$000101,\quad 000011,\quad 000110,\quad 001001,\quad 001010,$$
$$001100,\quad 010001,\quad 010010,\quad 010100,\quad 011000.$$

Каждая из этих строк начинается с нуля и состоит из $n=4$ нулей и $m=2$ единиц. Поэтому она определяется выборкой $n-1=3$ номеров мест нулей из $n+m-1=4+2-1=5$ номеров 1, 2, 3, 4, 5:

$$\{1, 2, 4\},\ \{1, 2, 3\},\ \{1, 2, 5\},\ \{1, 3, 4\},\ \{1, 3, 5\},$$
$$\{1, 4, 3\},\ \{2, 3, 4\},\ \{2, 3, 5\},\ \{2, 4, 5\},\ \{3, 4, 5\}$$

(считается, что нуль в начале каждой строки занимает место с номером 0).

Таким образом, задача о числе выборок с повторениями $m=2$ из $n \cdot m=8$ гостей $n=4$ пар сводится к задаче о числе выборок без повторений $n-1=4-1=3$ из $n+m-1=4+2-1=5$ номеров 1, 2, 3, 4, 5. В соответствии с этим число интересующих нас выборок выражается равенством

$$\binom{4+2-1}{4-1}=\binom{4+2-1}{3}=10.$$

Заметим, что в конечном счете каждому рассматриваемому отображению f была поставлена в соответствие выборка $Y=\{y_1, y_2, y_3\}$, элементы которой определяются равенствами

$$y_1=f(a_1)+1,\ y_2=f(a_1)+f(a_2)+2,\ y_3=f(a_1)+f(a_2)+ +f(a_3)+3, \text{ где } a_1=\mathscr{A},\ a_2=\mathscr{B},\ a_3=\mathscr{C},\ a_4=\mathscr{D}.$$

В частности, если $f(a_1)=f(a_2)=0$, $f(a_3)=f(a_4)=1$, то $y_1=1$, $y_2=2$, $y_3=4$; если $f(a_1)=f(a_2)=f(a_3)=0$, $f(a_4)=2$, то $y_1=1$, $y_2=2$, $y_3=3$.

Пример 3. *В детском наборе имеется по три одинаковых буквы с, а, т, е, l. Сколькими способами можно выбрать три из этих пятнадцати букв?*

Каждую такую выборку можно описать отображением f множества $A=\{c, a, m, e, l\}$ в множество $N=\{0, 1, 2, \ldots\}$ натуральных чисел: букв x выбирается $f(x)$ штук. Используя данный порядок букв c, a, m, e, l, составляющих множество A, эти отображения можно описать строками, в которых на месте с порядковым номером буквы x в множестве A расположено соответствующее этой букве значение $f(x)$ отображения f:

00111, 01011, 01101, 01110, 01002, 01020, 01200, ...

В частности, первая из этих строк описывает отображение f со значениями $f(c)=0$, $f(a)=0$, $f(m)=1$, $f(e)=1$, $f(l)=1$; вторая — со значениями $f(c)=0$, $f(a)=1$, $f(m)=0$, $f(e)=1$, $f(l)=1$.

Условимся натуральные числа 0, 1, 2, 3, ... обозначать также соответственно символами 0, 01, 011, 0,111, ... При этом условии рассматриваемые отображения f будут описываться также строками

00010101, 00100101, 00101001, 00101010,

00100011, 00100110, 00101100....

Каждая из этих строк начинается с нуля и состоит из $n=5$ нулей и $m=3$ единиц. Поэтому она определяется выборкой $n-1=5-1=4$ номеров мест нулей из $n+m-1=5+3-1=7$ номеров 1, 2, 3, 4, 5, 6, 7:

$\{1, 2, 4, 6\}$, $\{1, 3, 4, 6\}$, $\{1, 3, 5, 6\}$, $\{1, 3, 5, 7\}$,

$\{1, 3, 4, 5\}$, $\{1, 3, 4, 7\}$, $\{1, 3, 6, 7\}$, ...

(считается, что нуль в начале каждой строки занимает место с номером 0).

Таким образом, задача о числе выборок с повторениями $m=3$ из $n \cdot m=15$ изображений $n=5$ букв сводится к задаче о числе выборок без повторений $n-1=5-1=4$ из $n+m-1=5+3-1=7$ номеров 1, 2, 3, 4, 5, 6, 7. В соответствии с этим число интересующих нас выборок выражается равенством

$$\binom{5+3-1}{5-1}=\binom{5+3-1}{3}=35.$$

Заметим, что в конечном счете каждому рассматриваемому отображению f была поставлена в соответствие выборка

$$Y=\{y_1, y_2, y_3, y_4\},$$

элементы которой определяются равенствами $y_1=f(a_1)+1$, $y_2=f(a_1)+f(a_2)+2$, $y_3=f(a_1)+f(a_2)+f(a_3)+3$, $y_4=f(a_1)+f(a_2)+f(a_3)+f(a_4)+4$, где $a_1=c$, $a_2=a$, $a_3=m$, $a_4=e$, $a_5=l$.

В частности, если $f(a_1)=f(a_2)=0$, $f(a_3)=f(a_4)=f(a_5)=1$, то $y_1=1$, $y_2=2$, $y_3=4$, $y_4=6$; если $f(a_1)=0$, $f(a_2)=1$, $f(a_3)=0$, $f(a_4)=f(a_5)=1$, то $y_1=1$, $y_2=3$, $y_3=4$, $y_4=6$.

2.8. Формула для числа выборок с повторениями

Рассмотрим произвольные натуральные числа n, $m>0$ и множество A из n элементов. Каждое отображение f множества A в множество $N=\{0, 1, 2, \ldots\}$ натуральных чисел, сумма $\Sigma f(x)$ значений $f(x)$ которого равна m, назовем *выборкой с повторениями m из n элементов множества A*:

$$f: A \to N,\ \Sigma f(x)=m.$$

Докажем, что число всех таких выборок выражает

Формула для числа выборок с повторениями:

$$\binom{n+m-1}{n-1}=\binom{n+m-1}{m}=\frac{(n+m-1)}{m!\,(n-1)!}.$$

— Упорядочим множество A произвольной нумерацией его элементов: $A=\{a_1, \ldots, a_n\}$. При этом условии каждая выборка с повторениями f каких-либо m из n элементов множества A определяет выборку без повторений $Y=\{y_k: 1 \leqslant k \leqslant n-1\}$ возрастающих номеров $y_k=\sum_{1 \leqslant j \leqslant k} f(a_j)+k$ $(1 \leqslant k \leqslant n-1)$ из множества $B=\{1, \ldots, n+m-1\}$.

Различные отображения f определяют различные такие выборки Y. И каждая выборка $Y=\{y_k: 1 \leqslant k \leqslant n-1\}$ занумерованных в возрастающем порядке $n-1$ из $n+m-1$ элементов множества B определяется отображением $f: A \to N$ со значениями:

$$f(a_1)=y_1-1,$$

$$f(a_k)=y_k-y_{k-1}-1 \quad (2 \leqslant k \leqslant n-1),$$

$$f(a_n)=(n+m-1)-y_{n-1},$$

сумма которых, как нетрудно проверить, равна m:

$$\sum_{x\in A} f(x) = \sum_{1\leqslant k\leqslant n} f(a_k) = [y_{n-1} - (n-1)] + [(n+m-1) - y_{n-1}] = m.$$

Из сказанного вытекает, что число выборок с повторениями f произвольных m из n элементов множества A равно числу выборок без повторений Y произвольных $n-1$ из $n+m-1$ номеров множества B. Доказываемая формула для числа выборок с повторениями следует из доказанной формулы для числа выборок без повторений. +

Таким образом, *если имеется по m одинаковых предметов каждого из n различных типов, то m из этих $n\cdot m$ предметов можно выбрать* $\binom{n+m-1}{m}$ *способами.*

4.8. Задача о разложении числа на слагаемые

Формула для числа выборок с повторениями позволяет ответить на вопрос: *скольким способами натуральное число n можно представить в виде суммы k натуральных чисел, учитывая порядок слагаемых?*

Рассмотрим произвольные натуральные числа n, $k\geqslant 2$. Каждое семейство $x_1, \ldots, x_k$ натуральных чисел, в сумме равных n, условимся называть *положительным* решением уравнения

$$x_1+\ldots+x_k=n.$$

Если натуральные числа $x_1, \ldots, x_k$ строго положительны, то семейство $x_1, \ldots, x_k$ будем называть *строго положительным* решением уравнения $x_1+\ldots+x_k=n$.

Рассматриваемый вопрос о разложении натурального числа n на k натуральных слагаемых эквивалентен вопросу: *сколько существует положительных решений уравнения $x_1+\ldots+x_k=n$?*

Каждое положительное решение $x_1, \ldots, x_k$ является выборкой с повторениями f со значениями $f(1)=x_1, \ldots, f(k)=x_k$, составленной n элементами множества $A=\{1, \ldots, k\}$:

$$f(1)+\ldots+f(k)=n.$$

Число таких выборок f выражается равенствами

$$\binom{k+n-1}{k-1} = \binom{k+n-1}{n} = \frac{(k+n-1)!}{n!\,(k-1)!}.$$

Таким образом, существует $(k+n-1)!/n!(k-1)!$ положительных решений $x_1, \ldots, x_k$ уравнения $x_1+\ldots+x_k=n$ и столько же способов разложить натуральное число n на k натуральных слагаемых с учетом их порядка.

Пример. Число положительных решений x_1, x_2 уравнения

$$x_1+x_2=3$$

выражается равенствами

$$\binom{2+3-1}{1}=\binom{2+3-1}{3}=\frac{(2+3-1)!}{3!1!}=4.$$

Эти решения: 0, 3; 1, 2; 2, 1 и 3, 0.

Задача о числе строго положительных решений $x_1, \ldots, x_k$ уравнения $x_1+\ldots+x_k=n$ сводится к задаче о числе положительных решений $y_1, \ldots, y_k$ уравнения

$$y_1+\ldots+y_k=n-k.$$

Каждое строго положительное решение $x_1, \ldots, x_k$ уравнения $x_1+\ldots+x_k=n$ определяет положительное решение $y_1=x_1-1, \ldots, y_k=x_k-1$ уравнения $y_1+\ldots+y_k=n-k$. Различные решения $x_1, \ldots, x_k$ определяют различные решения $y_1, \ldots, y_k$, и каждое положительное решение $y_1, \ldots, y_k$ уравнения $y_1+\ldots+y_k=n-k$ определяется строго положительным решением $x_1=y_1+1, \ldots, x_k=y_k+1$ уравнения $x_1+\ldots+x_k=n$. Следовательно, число строго положительных решений $x_1, \ldots, x_k$ уравнения $x_1+\ldots+x_k=n$ равно числу

$$\binom{k+(n-k)-1}{k-1}=\binom{n-1}{k-1}$$

положительных решений $y_1, \ldots, y_k$ уравнения $y_1+\ldots+y_k=n-k$.

Пример. Число строго положительных решений x_1, x_2 уравнения

$$x_1+x_2=3$$

выражается равенством

$$\binom{3-1}{2-1}=2.$$

Эти решения: 1, 2 и 2, 1.

Часть II

ЗАДАЧИ

В этой части сначала рассматриваются задачи по комбинаторике, поясняющие содержание части I. Они дают представление о приложениях комбинаторики. Некоторые из них имеют определенное практическое значение.

Затем показывается, как использование различных комбинаторных правил позволяет решить трудную задачу о сумме степеней.

Глава 1.

ЗАДАЧИ ПО КОМБИНАТОРИКЕ

§ 1. ПРАВИЛА СЛОЖЕНИЯ И УМНОЖЕНИЯ

Для каждых конечных множеств A, B и для чисел элементов множеств A, B, $A+B$, $A\times B$ верны

Правило сложения:

$$n(A+B)=n(A)+n(B) \quad (AB=0).$$

Правило умножения:

$$n(A\times B)=n(A)\times n(B).$$

Аналогичные общие правила верны для нескольких множеств.

Используя правила сложения и умножения, можно решить следующие задачи.

Задача 1. *Доказать, что для каждой части A конечного множества B верно*

Правило вычитания:

$$n(B-A)=n(B)-n(A) \quad (A\subseteq B).$$

Решение. Каждая часть A конечного множества B не пересекается со своим дополнением $B-A$. По правилу сложения отсюда следует, что

$$n(B)=n((B-A)+A)=n(B-A)+n(A),$$

откуда вытекает доказываемое равенство.

Задача 2. *Доказать, что для каждых конечных множеств A и B верно*

Правило объединения:

$$n(A \cup B) = n(A) + n(B) - n(AB).$$

Решение. Из определений объединения и дополнения следует, что

$$A \cup B = A + (B - AB), \; A(B - AB) = 0, \; (AB = A \cap B)$$

для каждых конечных множеств A и B. По правилам сложения и вычитания отсюда вытекает, что

$$n(A \cup B) = n(A) + n(B - AB) = n(A) + n(B) - n(AB).$$

Задача 3. *Лекции по физике посещают 20 студентов, а лекции по астрономии — 30. Сколько студентов посещают лекции по физике или по астрономии, если:*

а) эти лекции происходят в одно и то же время;

б) эти лекции происходят в различное время и 10 студентов слушают оба курса?

Решение. Обозначим множество студентов, посещающих лекции по физике, буквой A, а лекции по астрономии,— буквой B.

а. В этом случае $AB = 0$ и, по правилу сложения,

$$n(A+B) = n(A) + n(B) = 20 + 30 = 50;$$

б. В этом случае $AB \neq 0$, и, по правилу объединения,

$$n(A \cup B) = n(A) + n(B) - n(A \cap B) = 20 + 30 - 10 = 40.$$

Задача 4. *Доказать, что для каждых конечных множеств A, B, и C верно равенство*

$$n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(A \cap C) - \\ - n(B \cap C) + n(A \cap B \cap C).$$

Решение. Используя правило объединения, последовательно получаем:

$$n(A \cup B \cup C) = n(A) + n(B \cup C) - n(A \cap (B \cup C)),$$

$$n(B \cup C) = n(B) + n(C) - n(B \cap C),$$

$$n(A \cap (B \cup C)) = n((A \cap B) \cup (A \cap C)) = n(A \cap B) + n(A \cap C) - \\ - n(A \cap B \cap C).$$

Из этих равенств следует доказываемое равенство.

Условимся называть равенство задачи 4 *правилом объединения для трех множеств.*

Задача 5. *В специализированном спортивном магазине, который торгует лыжами, лыжными ботинками и лыжными палками, за месяц было куплено 1000 пар лыж, 500 пар ботинок и 500 пар палок. При этом 400 пар лыж куплено вместе с ботинками, 300*

пар лыж — вместе с палками, 200 пар ботинок — вместе с палками, а 100 пар лыж — вместе с ботинками и палками.

Сколько покупателей посетило магазин?

Решение. Обозначим множество покупателей, приобретших лыжи, буквой A, ботинки — B, палки — C. Из условия задачи и правила объединения для трех множеств следует, что

$$n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(A \cap C) - n(B \cap C) + n(A \cap B \cap C) = 1000 + 500 + 500 - 400 - 300 - 200 + 100 = 1200.$$

Магазин посетило 1200 покупателей.

Задача 6. *На экзамене по математике было предложено три задачи: одна по алгебре, одна по геометрии, одна по тригонометрии. Из 1000 абитуриентов задачу по алгебре решили 800, по геометрии — 700, по тригонометрии — 600. При этом задачи по алгебре и геометрии решили 600 абитуриентов, по алгебре и тригонометрии — 500, по геометрии и тригонометрии — 400, а 300 абитуриентов решили все задачи.*

Сколько абитуриентов не решили ни одной задачи?

Решение. Обозначим множество всех абитуриентов буквой U, решивших задачу по алгебре — A, по геометрии — B, по тригонометрии — C. Из условия задачи и правила объединения для трех множеств следует, что

$$n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(A \cap C) - n(B \cap C) + n(A \cap B \cap C) = 800 + 700 + 600 - 600 - 500 - 400 + 300 = 900.$$

Хотя бы одну задачу решили 900 абитуриентов. По правилу вычитания

$$n(U - (A \cup B \cup C)) = n(U) - n(A \cup B \cup C) = 1000 - 900 = 100.$$

Ни одной задачи не решили 100 абитуриентов.

Задача 7. *Сколько всего существует строк длины 10, составленных из элементов множества $\{0, 1\}$?*

Решение. Из правила умножения следует, что

$$n(\{0, 1\} \times \{0, 1\} \times \ldots \times \{0, 1\}) = n(\{0, 1\}) \times \ldots \times n(\{0, 1\}) = 2^{10} = 1024.$$

1.1. Подбрасывание монеты

Производится опыт, состоящий из последовательности 10 испытаний, каждое из которых заключается в подбрасывании монеты один раз и имеет два возможных результата: монета падает цифрой вверх либо гербом. Результаты испытаний составляют результат опыта.

Задача 8. *Сколько всего существует возможных результатов опыта с десятью подбрасываниями монеты?*

Решение. Условимся результаты рассматриваемого опыта обозначать строками длины 10, составленными из элементов множества $\{0, 1\}$. Падение монеты цифрой вверх при первом подбрасывании обозначается элементом 0 на первом месте строки, гербом вверх — элементом 1. Падение монеты цифрой вверх при втором подбрасывании обозначается элементом 0 на втором месте строки, гербом вверх — элементом 1. Например, строка 00...0 обозначает исход опыта, при котором монета падает цифрой вверх при каждом подбрасывании, а строка 01 ... 1 — исход опыта, при котором монета падает цифрой вверх при первом подбрасывании и гербом вверх при всех остальных.

Различные результаты обозначаются различными строками, и каждая строка обозначает некоторый исход. Таким образом, число результатов рассматриваемого опыта равно числу строк длины 10, составленных из элементов множества $\{0, 1\}$, т. е. числу $2^{10}=1024$.

Задача 9. *Сколько всего существует строк длины 2, составленных из элементов множества $\{1, 2, 3, 4, 5, 6\}$?*

Решение. По правилу умножения

$$n(\{1, 2, 3, 4, 5, 6\}\times\{1, 2, 3, 4, 5, 6\})=n(\{1, 2, 3, 4, 5, 6\})\times$$
$$\times n(\{1, 2, 3, 4, 5, 6\})=6^2=36.$$

2.1. Подбрасывание кости

Производится опыт, состоящий из последовательности 2 испытаний, каждое из которых заключается в подбрасывании игральной кости один раз и имеет 6 возможных результатов: кость падает гранью 1 вверх, кость падает гранью 2 вверх, ..., кость падает гранью 6 вверх. Игральная кость представляет собой кубик, грани которого занумерованы цифрами 1, 2, 3, 4, 5, 6. Вместо «кость падает гранью 1 вверх» говорят «выпадает 1 очко», «кость падает гранью 2 вверх» — «выпадает 2 очка», ..., «кость падает гранью 6 вверх» — «выпадает 6 очков». Результаты испытаний составляют результат опыта.

Задача 10. *Сколько всего существует возможных результатов опыта с двумя подбрасываниями игральной кости?*

Решение. Условимся результаты рассматриваемого опыта обозначать строками длины 2, составленными из элементов множества $\{1, 2, 3, 4, 5, 6\}$. Выпадение 1 очка при первом подбрасывании обозначается цифрой 1 на первом месте строки, 2 очков — цифрой 2, ..., 6 очков — цифрой 6. Выпадение 1 очка при втором подбрасывании обозначается цифрой 1 на втором месте строки, 2 очков — цифрой 2, ..., 6 очков — цифрой 6. Например, строка 34 обозначает исход опыта, при котором выпадает 3 очка при первом подбрасывании и 4 очка — при втором, а строка 66 обозначает «двойную шестерку» — исход опыта, при котором выпадает 6 очков при первом подбрасывании и 6 очков — при втором.

Различные исходы обозначаются различными строками, и каждая строка обозначает некоторый исход. Таким образом, число исходов рассматриваемого опыта равно числу строк длины 2, составленных из элементов множества $\{1, 2, 3, 4, 5, 6\}$, т. е. числу $6^2=36$.

3.1. Выбор с возвращением

В урне находятся 5 шаров, отмеченных номерами 1, 2, 3, 4, 5.

Первое испытание заключается в том, что из урны вынимается наугад выбранный шар. Первое испытание имеет 5 возможных результатов: вынимается шар с номером 1, с номером 2, ..., с номером 5. Вынутый шар возвращается в урну. В этом случае после первого испытания в урне по-прежнему находятся 5 шаров с номерами 1, 2, 3, 4, 5.

Второе испытание заключается в том, что из урны второй раз вынимается наугад выбранный шар. Второе испытание имеет 5 возможных результатов: вынимается шар с номером 1, с номером 2, ..., с номером 5. Эти результаты не зависят от результатов первого испытания.

Результаты опыта, состоящего из первого и второго испытаний, составляются из результатов этих испытаний.

Задача 11. *Сколько всего существует возможных результатов опыта, состоящего в последовательном выборе с возвращением двух из пяти различных шаров?*

Решение. Условимся результаты рассматриваемого опыта обозначать строками u_1u_2 длины 2, составленными из элементов u_1 и u_2 множества $\{1, 2, 3, 4, 5\}$. Строка u_1u_2 обозначает результат «первым вынут шар с номером u_1, вторым вынут шар с номером u_2». Различные результаты обозначаются различными строками, и каждая строка обозначает некоторый результат. Таким образом, число результатов рассматриваемого опыта равно числу всех строк длины 2, составленных из элементов множества 1, 2, 3, 4, 5, т. е. числу $5^2=25$.

4.1. Выбор без возвращения

В урне находятся 5 шаров, отмеченных номерами 1, 2, 3, 4, 5.

Первое испытание заключается в том, что из урны вынимается наугад выбранный шар. Первое испытание имеет 5 результатов: вынимается шар с номером 1, с номером 2, ..., с номером 5. Вынутый шар не возвращается в урну. В этом случае в урне остаются 4 шара. Номера этих шаров определяются номером вынутого шара. Например, если вынимается шар с номером 1, то остаются шары с номерами 2, 3, 4, 5.

Второе испытание заключается в том, что из урны второй раз вынимается наугад выбранный шар. Второе испытание имеет 4

возможных результата. Эти результаты зависят от результата первого испытания. Например, если при первом испытании вынимается шар с номером 1, то при втором может быть вынут шар с номером 2, либо 3, либо 4, либо 5.

Результаты опыта, состоящего из первого и второго испытаний, составляются из результатов этих испытаний.

Задача 12. *Сколько всего существует возможных результатов опыта, состоящего в последовательном выборе без возвращения двух из пяти различных шаров?*

Решение. Условимся результаты рассматриваемого опыта обозначать строками u_1u_2 длины 2, составленными из неравных элементов u_1 и u_2 множества $\{1, 2, 3, 4, 5\}$. Строка u_1u_2 $(u_1 \neq u_2)$ обозначает результат «первым вынут шар с номером u_1, вторым вынут шар с номером u_2». Различные результаты обозначаются различными такими строками, и каждая такая строка обозначает некоторый результат. Таким образом, число результатов рассматриваемого опыта равно числу строк длины 2, составленных из неравных элементов множества $\{1, 2, 3, 4, 5\}$. По правилу вычитания, это число равно числу всех строк длины 2, составленных из элементов множества $\{1, 2, 3, 4, 5\}$, минус число строк длины 2, составленных из равных элементов множества $\{1, 2, 3, 4, 5\}$, т. е. числу $25-5=20$.

5.1. Примеры из генетики

Рассмотрим несколько простых задач, имеющих отношение к генетике. Разъяснение употребляемых терминов можно найти в книге Н. П. Дубинина «Горизонты генетики» (М., 1970).

Пример 1. Гаметы. Простейшая модель образования половых клеток в организме (*гамет*), связана с распределением каждой *хромосомы* каждой из имеющихся в *диплоидном* наборе n пар в одну из двух образующихся гамет. При этом распределении каждая хромосома каждой пары может попасть в каждую гамету независимо от хромосом других пар.

Хромосомы каждой пары *гомологичны*, а хромосомы каждых различных пар — нет.

Гаплоидный набор n негомологичных хромосом, попавших в гамету, можно описать строкой длины n, на $1, \ldots, n$ месте которой расположен номер 1 или 2 попавшей в гамету хромосомы $1, \ldots, n$ пары. Поэтому число всех таких наборов n хромосом равно числу всех таких строк длины n номеров 1, 2.

Предложение. *Существует ровно 2^n строк длины n, составленных из двух элементов.*

Доказательство. Рассмотрим произвольное множество из двух элементов a и b.

Строк длины $n=1$ таких элементов ровно две:

$$2^1=2.$$

Для каждого номера n множество всех строк длины $n+1$ лементов a, b разбивается на два непересекающихся множества A и B, состоящие соответственно из строк с элементами a и b на $n+1$ месте. Число строк в каждом из множеств A и B равно числу строк длины n элементов a, b. Если предположить, что это число равно 2^n, то число всех рассматриваемых строк длины $n+1$ будет, по правилу сложения, выражаться равенствами

$$n(A+B)=n(A)+n(B)=2^n+2^n=2\cdot 2^n=2^{n+1}.$$

По принципу индукции из сказанного следует, что доказываемое предложение верно для каждого натурального числа $n\geqslant 1$.

Таким образом, в рассматриваемой модели *существует ровно 2^n гаплоидных наборов n хромосом*, каждый из которых может попасть в образующуюся гамету.

Знаменитая дрозофила меланогастер характеризуется диплоидным набором $n=4$ пар хромосом. Число гаплоидных наборов хромосом для нее равно

$$2^4=16.$$

Человек характеризуется диплоидным набором $n=23$ пар хромосом. Число гаплоидных наборов хромосом для него равно

$$2^{23}=8\,388\,608.$$

Остальные примеры связаны с одним из классических опытов Грегора Менделя с горохом.

Пример 2. Генотип. Пусть *гены* A, B, C определяют соответственно круглые семена, желтые семядоли, красные цветы растения. А *аллельные* им гены a, b, c — морщинистые семена, зеленые семядоли, белые цветы.

Аллельные гены занимают одинаковые места (*локусы*) в гомологичных хромосомах. Поэтому диплоидный набор хромосом характеризуется числами z_1, z_2, z_3 содержащихся в нем соответственно генов A, B, C. Каждое из чисел z_1, z_2, z_3 может равняться каждому из чисел 2, 1, 0. Строка $z_1z_2z_3$ описывает *генотип* растения по рассматриваемым трем признакам.

В частности, строка 222 описывает *чистый* генотип, образованный парами AA, BB, CC одинаковых аллельных генов, строка 000 — чистый генотип, образованный парами aa, bb, cc одинаковых аллельных генов, строка 111 — *смешанный* генотип, образованный парами Aa, Bb, Cc различных аллельных генов.

Таким образом, в рассматриваемой модели число различных генотипов равно числу всех строк длины $n=3$ чисел 2, 1, 0. Последовательно применяя правило умножения, убеждаемся в том, что это число выражается равенствами

$$(3\cdot 3)3=3^2\cdot 3=27.$$

Строки, описывающие генотипы, нетрудно выписать, разбив их на множества строк с нулями на данных местах, а эти множества разбив на классы по числу нулей в строке:

Класс	0	1			2			3
Множество	Φ_0	Φ_1	Φ_2	Φ_3	Φ_{12}	Φ_{13}	Φ_{23}	Φ_{123}
Строка	222 221 212 122 211 121 112 111	022 021 012 011	202 201 102 101	220 210 120 110	002 001	020 010	200 100	000

Пример 3. Фенотип. В примере 2 гены A, B, C являются *доминантными,* а гены a, b, c — *рецессивными.* Совокупность признаков (или *фенотип* растения) определяется доминантными генами A, B, C. Если в генотипе есть хотя бы один ген A, то растение имеет круглые семена, если нет,— морщинистые. Если в генотипе есть хотя бы один ген B, то растение имеет желтые семядоли, если нет,— зеленые. Если в генотипе есть хотя бы один ген C, то растение имеет красные цветы, если нет,— белые.

Различные генотипы могут определять одинаковые фенотипы. Условимся каждый рассматриваемый фенотип сокращенно записывать начальными буквами слов, описывающих форму семян, цвет семядолей, окраску цветов. В частности, буквы К, Ж, К обозначают растение с круглыми семенами, желтыми семядолями, красными цветами, буквы М, З, Б — растение с морщинистыми семенами, зелеными семядолями, белыми цветами.

Используя таблицу примера 2, можно описать соответствие между рассматриваемыми генотипами и фенотипами, разбив фенотипы на классы по числу рецессивных признаков:

Класс	0	1			2			3
Генотип	Φ_0	Φ_1	Φ_2	Φ_3	Φ_{12}	Φ_{13}	Φ_{23}	Φ_{123}
Фенотип	КЖК	МЖК	КЗК	КЖБ	МЗК	МЖБ	КЗБ	МЗБ

Общее число фенотипов выражается равенством

$$1+3+3+1=8.$$

В соответствии с правилом сложения, сумма чисел генотипов в множествах Φ_0, Φ_1, Φ_2, Φ_3, Φ_{12}, Φ_{13}, Φ_{23}, Φ_{123} равна числу всех рассматриваемых генотипов:

$$8+3\cdot4+3\cdot2+1=27.$$

Пример 4. Гаметы. Образование гамет у рассматриваемых растений связано с независимым распределением каждой аллели каждой из пар в одну из двух образующихся гамет. При этом распределении каждая аллель каждой пары может попасть в каждую гамету.

Гаплоидный набор хромосом в образующихся гаметах характеризуется числами x_1, x_2, x_3 и y_1, y_2, y_3 генов A, B, C, содержащихся в одной и другой гамете соответственно. Каждое из этих чисел может равняться каждому из чисел 1, 0. Строки $x_1x_2x_3$ и $y_1y_2y_3$ описывают генотипы образующихся гамет. Число таких генотипов равно $2^3=8$.

В частности, строка 111 описывает генотип ABC гаметы, образованный генами A, B, C, строка 000 — генотип abc гаметы, образованный генами a, b, c.

Генотипы $x_1x_2x_3$, $y_1y_2y_3$ образующихся гамет связаны с генотипом $z_1z_2z_3$ порождающего их организма уравнениями

$$(*)\qquad x_1+y_1=z_1, \quad x_2+y_2=z_2, \quad x_3+y_3=z_3.$$

Число пар строк $x_1x_2x_3$ и $y_1y_2y_3$, образующих решение этой системы уравнений для каждой строки $z_1z_2z_3$, описывает число пар генотипов гамет, образуемых растением, имеющим генотип $z_1z_2z_3$.

Обозначим символом $n(z)$ число допустимых решений системы (*), описывающих генотипы гамет, а символом $s(z)$ — число единиц в строке $z=z_1z_2z_3$ правых частей этой системы.

Теорема. $n(z)=2^{s(z)}$.

Доказательство. Число $n(z)$ решений $x_1x_2x_3$ и $y_1y_2y_3$ системы уравнений (*) равно числу троек, составленных из решений x_1 и y_1, x_2 и y_2, x_3 и y_3 соответственно 1, 2, 3-му уравнениям этой системы.

Решения x_1 и y_1 первого уравнения описываются таблицей

x_1	y_1	z_1
1	1	2
1	0	1
0	1	1
0	0	0

Эта таблица показывает, что для числа $n_1(z)$ таких решений и числа $s_1(z)$ единиц на первом месте строки $z=z_1z_2z_3$ верно равенство

$$n_1(z) = 2^{s_1(z)}.$$

Аналогичные равенства верны для чисел $n_2(z)$, $n_3(z)$ решений второго и третьего уравнений (*) и чисел $s_2(z)$, $s_3(z)$ единиц на втором, третьем местах строки $z=z_1z_2z_3$:

$$n_2(z)=2^{s_2(z)},\quad n_3(z)=2^{s_3(z)}.$$

По принципу умножения из сказанного следует, что

$$n(z)=n_1(z)\,n_2(z)\,n_3(z)=2^{s_1(z)}\cdot 2^{s_2(z)}\cdot 2^{s_3(z)}=2^{s_1(z)+s_2(z)+s_3(z)}=2^{s(z)}.$$

Теорема доказана.

Из этой теоремы, в частности, следует, что каждое растение, имеющее чистый генотип, образованный парами одинаковых генов, порождает гаметы только одного типа:

$$n(222)=n(220)=n(202)=n(022)=n(200)=n(020)=$$
$$=n(002)=n(000)=2^0=1.$$

А каждое растение, имеющее смешанный генотип, хотя бы одна пара аллелей которого образована различными генами, порождает гаметы различных типов, в частности,

$$n(111)=2^{s(111)}=2^3=8,$$

и, следовательно, гибрид, имеющий генотип 111, может производить гаметы, имеющие любой из возможных генотипов.

Пример 5. Зигота. При слиянии двух гамет образуется зародышевая клетка (*зигота*). Гаплоидные наборы хромосом сливающихся гамет образуют диплоидный набор хромосом зиготы. Генотип $z_1z_2z_3$ зиготы связан с генотипами $x_1x_2x_3$ и $y_1y_2y_3$ образующих ее гамет уравнениями

$$(*)\quad x_1+y_1=z_1,\ x_2+y_2=z_2,\ x_3+y_3=z_3.$$

Доказанная теорема позволяет подсчитывать число пар $x_1x_2x_3$ и $y_1y_2y_3$ генотипов гамет, образующих зиготу с генотипом $z=z_1z_2z_3$. Зигота с чистым генотипом образуется парой гамет только одного типа, зигота со смешанным генотипом — парами гамет различных типов.

Совпадение систем уравнений (*), описывающих образование гамет и зиготы, выражает способность растений к воспроизведению. Растение с чистым генотипом воспроизводит себя с помощью гамет одного типа. Растение со смешанным генотипом с помощью своих гамет различных типов может производить растения, имеющие другие генотипы.

§ 2. ПЕРЕСТАНОВКИ

Рассмотрим множества A и B, состоящие каждое из n элементов. Подстановкой вместо A множества B называется взаимно однозначное отображение множества A на множество B.

Существует ровно $n!=1\cdot 2\cdot\ldots\cdot n$ подстановок вместо A множества B:

$$0!=1,\ 1!=1,\ 2!=1\cdot 2=2,\ 3!=1\cdot 2\cdot 3=6,$$
$$n!=1\cdot 2\cdot \ldots \cdot n \quad (n\geqslant 3).$$

Это предложение позволяет решить классическую задачу о числе перестановок.

Задача. *Сколькими способами можно переставить n различных предметов, расположенных на n различных местах?*

Ответ: $n!=1\cdot 2\cdot\ldots\cdot n$ способами.

К задаче о числе перестановок сводятся следующие задачи.

Задача 1. *Сколькими способами можно расположить в ряд на книжной полке 5 различных книг?*

Решение. Число всех таких расположений равно числу всех перестановок множества, составленного из 5 элементов, т. е. числу

$$5!=1\cdot 2\cdot 3\cdot 4\cdot 5=120.$$

Задача 2. *Сколько всего сигналов можно составить, меняя порядок семи различных флагов: красного, синего, зеленого, желтого, коричневого, черного и белого?*

Решение. Число всех таких сигналов равно числу всех перестановок множества, составленного из 7 элементов, т. е. числу

$$7!=1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7=5040.$$

Задача 3. *Сколькими способами можно расставить на шахматной доске 8 ладей так, чтобы никакие 2 из них не были на одной горизонтали или вертикали?*

Решение. Рассмотрим множество $A=\{1, 2, 3, 4, 5, 6, 7, 8\}$ номеров горизонталей и множество $B=\{a, b, c, d, e, f, g, h\}$ букв, обозначающих вертикали шахматной доски.

По условию задачи на каждой горизонтали с номером x должна стоять ладья. Пусть $s(x)$ — буква, обозначающая вертикаль, на которой находится эта ладья. Кроме того, по условию задачи на одной вертикали не может быть 2 ладьи или больше. Значит, буквы $s(x)$ различны для различных номеров x. Следовательно, каждое расположение ладей, удовлетворяющее условию задачи, описывается подстановкой s вместо A множества B. Например, расположения ладей по диагоналям доски описываются подстановками

$$\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7 & 8\\ a & b & c & d & e & f & g & h\end{pmatrix},\ \begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7 & 8\\ h & g & f & e & d & c & b & a\end{pmatrix}.$$

Искомое число расположений ладей на доске равно числу всех подстановок вместо A множества B, т. е. числу

$$8!=1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7\cdot 8=40\,320.$$

Задача 4. *Сколькими способами можно распределить 10 различных писем по 10 различным конвертам?*

Решение. Число всех возможных распределений 10 различных писем по 10 различным конвертам равно числу всех перестановок множества, составленного из 10 элементов, т. е. числу

$$10!=1\cdot 2\cdot\ldots\cdot 10=3\,628\,800.$$

Задача 5. *Алхимик знает, что для того, чтобы приготовить элексир жизни, нужно смешать в определенных пропорциях имеющиеся у него 20 различных жидкостей. Ему не известно только, в каком порядке следует смешивать эти жидкости. Сколько всего существует таких порядков?*

Решение. Число всех возможных порядков смешивания 20 различных жидкостей равно числу всех перестановок множества, составленного из 20 элементов, т. е. числу 20!. Как нетрудно проверить, число 20! больше числа 10^{16}. Если на приготовление каждой смеси алхимик будет тратить одну минуту, то на проверку всех возможных смесей ему потребуется больше 10^{10} лет. Этот срок значительно превышает предполагаемый возраст нашей солнечной системы.

Разложение молекул белка приводит к 20 стандартным аминокислотам. Молекулы нуклеиновых кислот, содержащихся в хромосомах, составляются из 4 нуклеотидов: Аденина, Тимина, Гуанина, Цитозина. Из этих 4 нуклеотидов можно составить 20 различных троек:

ААА ААТ ААГ ААЦ АТТ АТГ АТЦ АГГ АГЦ АЦЦ
ТТТ ТТГ ТТЦ ТГГ ТГЦ ТЦЦ ГГГ ГГЦ ГЦЦ ЦЦЦ

(в этих тройках не учитывается порядок нуклеотидов, и некоторые из нуклеотидов повторяются).

Можно предположить, что между 20 аминокислотами и 20 тройками нуклеотидов существует некоторое соответствие — *генетический код.*

Задача 6. *Сколькими способами каждой из 20 аминокислот можно поставить в соответствие взаимно однозначным образом некоторую из 20 троек нуклеотидов?*

Решение. Число всех возможных взаимооднозначных соответствий между 20 аминокислотами и 20 тройками нуклеотидов равно числу всех подстановок вместо множества A из 20 элементов множества B из 20 элементов, т. е. числу 20!.

Как было отмечено, это число больше 10^{16}. Поэтому было бы безнадежно пытаться раскрыть генетический код, сопоставляя аминокислоты и тройки нуклеотидов наугад.

§ 3. ВЫБОРКИ

Выборкой (или сочетанием) m из n элементов множества A, составленного из n элементов, называется часть X множества A, составленная из m элементов ($m \leqslant n$). Существует ровно

$$\binom{n}{m} = \frac{n!}{m!\,(n-m)!}$$

выборок m из n элементов множества A.

Это предположение позволяет решить классическую задачу о числе выборок.

Задача. *Сколькими способами можно выбрать m из n различных предметов?*

Ответ: $\binom{n}{m}$ способами.

Для чисел $\binom{n}{m}$ верны

Правило симметрии:

$$\binom{n}{m} = \binom{n}{n-m},$$

Правило Паскаля:

$$\binom{n+1}{m} = \binom{n}{m} + \binom{n}{m-1}.$$

К задаче о числе выборок сводятся следующие задачи.

Задача 1. *Сколькими способами можно выбрать 3 из 5 различных карт?*

Ответ: $\binom{5}{3} = \frac{5!}{3!2!} = \frac{1\cdot2\cdot3\cdot4\cdot5}{1\cdot2\cdot3\cdot1\cdot2} = 10$ способами.

Задача 2. *Сколькими способами можно выбрать 13 из 52 различных карт?*

Ответ: $\binom{52}{13} = \frac{52!}{13!39!} = 635\,013\,559\,600$ способами.

Задача 3. *Сколькими способами можно составить комиссию в составе 3 человек, выбирая их из 4 супружеских пар, если:*

а) в комиссию могут входить любые 3 из 8 человек;

б) в комиссию не могут входить члены одной семьи?

Решение. а) Если в комиссию могут входить любые 3 из 8 человек, то существует всего

$$\binom{8}{3} = \frac{8!}{3!5!} = \frac{1\cdot2\cdot3\cdot4\cdot5\cdot6\cdot7\cdot8}{1\cdot2\cdot3\cdot1\cdot2\cdot3\cdot4\cdot5} = 56$$

таких комиссий.

б) Если в комиссию не могут входить члены одной семьи, то в ней должны быть представлены 3 из 4 семей. Эти семьи можно выбрать

$$\binom{4}{3} = \frac{4!}{3!1!} = \frac{1\cdot2\cdot3\cdot4}{1\cdot2\cdot3\cdot1} = 4$$

способами. После этого в каждой из них можно 2 способами выбрать представителя — мужа или жену. По правилу умножения отсюда следует, что можно составить всего

$$\binom{4}{3} 2\cdot2\cdot2 = 32$$

такие комиссии.

Задача 4. *Сколькими способами можно выбрать 2 из 5 различных шаров?*

Ответ: $\binom{5}{2}=\frac{5!}{2!3!}=\frac{1\cdot2\cdot3\cdot4\cdot5}{1\cdot2\cdot1\cdot2\cdot3}=10$ способами.

Рассмотрим m частиц, распределенных по n различным областям пространства. Предположим, что частицы не отличаются друг от друга и что в каждой области пространства может находиться не более одной частицы (и, следовательно, $m\leqslant n$). Состояние рассматриваемой системы определяется указанием областей пространства, в которых находятся частицы.

Условимся описанную модель называть моделью Ферми — Дирака.

Задача 5. *Сколько всего существует возможных состояний системы в модели Ферми — Дирака?*

Решение. Число состояний системы в модели Ферми — Дирака с m частицами и n областями пространства равно числу выборок m из n элементов, т. е. числу $\binom{n}{m}$. В частности, если $m=10$ и $n=100$, то

$$\binom{100}{10}=\frac{100!}{10!\,90!}=\frac{91\cdot92\cdot93\cdot94\cdot95\cdot96\cdot97\cdot98\cdot99\cdot100}{1\cdot2\cdot3\cdot4\cdot5\cdot6\cdot7\cdot8\cdot9\cdot10}=17\,310\,309\,456\,440.$$

§ 4. РАЗМЕЩЕНИЯ

Размещением по m местам множества B m из n элементов множества A называется взаимно однозначное отображение на множество B части X множества A, составленной из m элементов ($m\leqslant n$). Существует ровно

$$n^{(m)}=\frac{n!}{(n-m)!}$$

размещений по m местам множества B m из n элементов множества A.

Это предложение позволяет решить классическую задачу о числе размещений.

Задача. *Сколькими способами можно выбрать и разместить по m различным местам m из n различных предметов?*

Ответ: $n^{(m)}$ способами.

Если $m=n$, то задача о числе размещений сводится к задаче о числе перестановок.

К задаче о числе размещений сводятся следующие задачи.

Задача 1. *Сколькими способами можно выбрать и разместить в ряд на книжной полке 3 из 5 различных книг?*

Ответ: $5^{(3)}=5\cdot4\cdot3=60$ способами.

Задача 2. *Сколькими способами можно выбрать и разместить в ряд 2 из 5 различных шаров?*

Ответ: $5^{(2)}=5\cdot4=20$ способами.

Задача 3. *Сколько всего существует телефонных номеров, состоящих из 5 различных цифр?*

Решение. Число таких номеров равно числу размещений по 5 местам 5 из 10 цифр 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, т. е. числу

$$10^{(5)}=10\cdot9\cdot8\cdot7\cdot6=30\,240.$$

Задача 4. *Сколькими способами можно обозначить треугольник, как обычно отмечая его вершины большими буквами латинского алфавита?*

Ответ: $26^{(3)}=26\cdot25\cdot24=15\,600$ способами.

Задача 5. *Сколько различных билетов с указанием станции отправления и станции назначения можно отпечатать для железной дороги, на которой 50 станций?*

Ответ: $50^{(2)}=50\cdot49=2450$ способами.

§ 5. ФОРМУЛА НЬЮТОНА ДЛЯ БИНОМА

Для произвольных чисел x, y и натурального числа n верно равенство

$$(x+y)^n=\sum_{m=0}^{n}\binom{n}{m}x^m y^{n-m},$$

которое можно записать также в следующем развернутом виде:

$$(x+y)^n=\binom{n}{0}x^0y^{n-0}+\binom{n}{1}x^1y^{n-1}+\ldots+\binom{n}{m}x^my^{n-m}+\ldots$$
$$\ldots+\binom{n}{m}x^ny^{n-n}\qquad(0\leqslant m\leqslant n).$$

Используя это равенство, называемое формулой Ньютона для бинома, можно решить следующие задачи.

Задача 1. *Доказать, что для каждого натурального числа n верно равенство*

$$\sum_{m=0}^{n}\binom{n}{m}=2^n.$$

Решение. Используя формулу Ньютона для чисел $x=y=1$, получаем

$$2^n=(1+1)^n=\sum_{m=0}^{n}\binom{n}{m}1^m\cdot1^{n-m}=\sum_{m=0}^{n}\binom{n}{m}.$$

Задача 2. *Сколько всего существует частей множества, составленного из n элементов?*

Решение. Рассмотрим множество A, составленное из n элементов, и класс $\mathscr{P}(A)$ всех частей множества A. Задача заключается в определении числа элементов множества $\mathscr{P}(A)$. Для каждого номера $m\leqslant n$ рассмотрим класс $\mathscr{P}_m(A)$ всех частей множества A, составленных из m элементов. Ясно, что множества $\mathscr{P}_m(A)$ попарно

не пересекаются и их объединение равно множеству $\mathscr{P}(A)$. Например, если $A=\{0, 1\}$, то

$$\mathscr{P}_0(A)=\{0\},\ \mathscr{P}_1(A)=\{\{0\},\ \{1\}\},\ \mathscr{P}_2(A)=\{A\},$$
$$\mathscr{P}(A)=\{O,\ \{0\},\ \{1\}A\}.$$

По правилу объединения, формуле для числа выборок и равенству задачи 1 из сказанного следует, что

$$n(\mathscr{P}(A)) = \sum_{m=0}^{n} n(\mathscr{P}_m(A)) = \sum_{m=0}^{n} \binom{n}{m} = 2^n.$$

Таким образом, существует ровно 2^n частей множества, составленного из n элементов.

Задача 3. *Сколькими способами можно выбрать некоторые из n различных предметов?*

Решение. Число способов, которыми можно выбрать 0, 1, ..., m, ..., n предметов из n различных предметов, равно числу всех частей множества, составленного из n элементов, т. е. числу 2^n.

Задача 4. *Доказать, что для каждого натурального числа $n>0$ верно равенство*

$$\sum_{m=0}^{n} (-1)^m \binom{n}{m} = 0.$$

Решение. Используя формулу Ньютона для чисел $x=-1$ и $y=1$, получаем:

$$0=(-1+1)^n = \sum_{m=0}^{n} \binom{n}{m} (-1)^m\, 1^{n-m} = \sum_{m=0}^{n} \binom{n}{m} (-1)^m.$$

Задача 5. *Доказать, что для каждого натурального числа $n>0$ верны равенства*

$$\sum_{0\leqslant 2m\leqslant n} \binom{n}{2m} = \sum_{0\leqslant 2m+1\leqslant n} \binom{n}{2m+1} = 2^{n-1}.$$

Решение. Обозначим рассматриваемые суммы четных и нечетных биномиальных коэффициентов буквами a и b соответственно:

$$a = \sum_{0\leqslant 2m\leqslant n} \binom{n}{2m},\quad b = \sum_{0\leqslant 2m+1\leqslant n} \binom{n}{2m+1}.$$

Из равенств задач 1 и 4 следует, что

$$a+b=2^n,\ a-b=0,$$

и значит,

$$a=b=2^{-1}\cdot 2^n=2^{n-1}.$$

Задача 6. *Сколько всего существует составленных из четного числа элементов частей множества из n элементов?*

Ответ: 2^{n-1} $(n>0)$. Столько же существует составленных из нечетного числа элементов частей рассматриваемого множества.

Задача 7. *Сколькими способами можно выбрать четное число предметов из n различных предметов?*

Ответ: 2^{n-1} $(n>0)$. Столькими же способами можно выбрать нечетное число предметов.

Задача 8. *Какие из биномиальных коэффициентов* $\binom{n}{0}, \ldots, \binom{n}{m}, \ldots, \binom{n}{n}$ *являются наибольшими?*

Решение. Сравним коэффициенты для каждых соседних номеров m и $m+1$. Как нетрудно проверить, для каждых натурального числа n и номера $m<n$ верны равенства

$$\binom{n}{m+1}\Big/\binom{n}{m} = \frac{n!}{(m+1)!\,(n-m-1)!}\Bigg/\frac{n!}{m!\,(n-m)!} = \frac{n-m}{m+1}.$$

Рассмотрим для номера m следующие три возможные случая:

1. Если $m+1 \leqslant \frac{n}{2}$, то $2m+2 \leqslant n$, $m+2 \leqslant n-m$, $m+1 < n-m$ и

$$(1) \qquad \binom{n}{m+1}\Big/\binom{n}{m} = \frac{n-m}{m+1} > 1 \qquad \left(m+1 \leqslant \frac{n}{2}\right).$$

2. Если $m \geqslant \frac{n}{2}$, то $2m \geqslant n$, $m \geqslant n-m$ и $m+1 > n-m$ и

$$(2) \qquad \binom{n}{m+1}\Big/\binom{n}{m} = \frac{n-m}{m+1} < 1 \qquad \left(m \geqslant \frac{n}{2}\right).$$

3. Если $m < \frac{n}{2} < m+1$, то $2m < n < 2m+2$, $n = 2m+1$, $m+1 = n-m$ и

$$(3) \qquad \binom{n}{m+1}\Big/\binom{n}{m} = \frac{n-m}{m+1} = 1 \qquad \left(m < \frac{n}{2} < m+1\right).$$

Рассмотрим для номера n два возможные случая:

1. Предположим, что число n четное, и рассмотрим номер

$$\hat{m} = \frac{n}{2}.$$

Последовательно используя соотношения (1), легко убедиться в том, что для каждого номера $m < \hat{m}$ верно неравенство

$$\binom{n}{m} < \binom{n}{\hat{m}}.$$

Аналогично, последовательно используя соотношения (2), легко убедиться в том, что это неравенство верно и для каждого номера $m > \hat{m}$.

Например, для номера $m=\hat{m}-1$ из соотношений 1 следует, что

$$\binom{n}{\hat{m}}\Big/\binom{n}{m}=\binom{n}{m+1}\Big/\binom{n}{m}>1 \qquad \left(m+1=\hat{m}=\frac{n}{2}\right).$$

А для номера $m=\hat{m}+1$ из соотношений (2) следует, что

$$\binom{n}{m}\Big/\binom{n}{\hat{m}}=\binom{n}{\hat{m}+1}\Big/\binom{n}{\hat{m}}<1 \qquad \left(m=\hat{m}+1>\frac{n}{2}\right).$$

Таким образом, если число n четное, то наибольшим является коэффициент с номером $\hat{m}=\frac{n}{2}$. Все остальные коэффициенты строго меньше его.

2. Предположим теперь, что число n нечетное, и рассмотрим номер

$$\hat{m}=\frac{n-1}{2}.$$

Последовательно используя соотношения (1), легко убедиться в том, что для каждого номера $m<\hat{m}$ верно неравенство

$$\binom{n}{m}<\binom{n}{\hat{m}}.$$

Аналогично, последовательно используя соотношения (2), легко убедиться в том, что это неравенство верно и для каждого номера $m>\hat{m}+1$.

Так как $\hat{m}=\frac{(n-1)}{2}<\frac{n}{2}<\frac{n+1}{2}=\hat{m}+1$, то из соотношений (3) следует, что

$$\binom{n}{\hat{m}}=\binom{n}{\hat{m}+1}.$$

Таким образом, если число n нечетное, то наибольшими являются коэффициенты с номерами $\hat{m}=\frac{n-1}{2}$ и $\hat{m}+1=\frac{n+1}{2}$. Все остальные коэффициенты строго меньше их.

Задача 9. *Доказать, что для каждых натуральных чисел l и n верно равенство*

$$\sum_{0\leqslant m\leqslant l}\binom{n+m}{m}=\binom{n+l+1}{l}.$$

Решение. Заметим, что

$$\binom{n}{0}=\binom{n+1}{0}=1.$$

Заменив слагаемое с номером $m=0$ и последовательно используя правило Паскаля, получаем:

$$\sum_{0\leqslant m\leqslant l}\binom{n+m}{m}=\left[\binom{n+1}{0}+\binom{n+1}{1}\right]+\binom{n+2}{2}+\binom{n+3}{3}+\ldots+\binom{n+l-1}{l-1}+\binom{n+l}{l}=$$
$$=\left[\binom{n+2}{1}+\binom{n+2}{2}\right]+\binom{n+3}{3}+\ldots+\binom{n+l-1}{l-1}+\binom{n+l}{l}=$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$=\left[\binom{n+l}{l-1}+\binom{n+l}{l}\right]=$$
$$=\binom{n+l+1}{l}$$

Задача 10. *Доказать, что для каждых строго положительного числа $b<1$ и натурального числа $n>1$ верно неравенство*

$$b^n<\frac{1}{n\,(1-b)}.$$

Решение. Используя формулу Ньютона для чисел $x=b^{-1}-1>0$ и $y=1$, получаем:

$$b^{-n}=\left(b^{-1}\right)^n=(x+1)^n=\sum_{0\leqslant m\leqslant n}\binom{n}{m}x^m\geqslant 1+nx+\frac{n\,(n-1)}{2}x^2>$$
$$>nx=n(b^{-1}-1).$$

Следовательно,

$$b^n<1/n(b^{-1}-1)=b/n(1-b)<1/n(1-b),$$

что и требовалось доказать.

Пример. $\left(1-\frac{1}{1000}\right)^{1000}<\frac{1}{1000\,\frac{1}{100}}=\frac{1}{10}.$

Замечание. Неравенство задачи 10 может быть получено из неравенства Бернулли для $c=1/b$ $(0<b<1)$:

$$\frac{1}{b^n}>1+n\left(\frac{1}{b}-1\right)>n\left(\frac{1}{b}-1\right)>n\,(1-b)\,\frac{1}{b}>n\,(1-b).$$

§ 6. РАЗБИЕНИЯ

Каждое семейство попарно не пересекающихся множеств $A_1,\ldots,A_k$ из $n_1,\ldots,n_k$ элементов, в сумме составляющих множество A из n элементов, называется $n_1,\ldots,n_k$-разбиением множества A. Существует ровно

$$\binom{n}{n_1,\ldots,n_k}\frac{n!}{n_1!\ldots,n_k!}$$

$n_1,\ldots,n_k$-разбиений множества из n элементов $(n_1+n_2+\ldots+n_k=$ $=n)$.

Это предложение позволяет решить задачу о числе разбиений.

Задача. *Сколькими способами можно разбить n различных предметов на k групп $1, \ldots, k$ по $n_1, \ldots, n_k$ предметов?*

Ответ: $\binom{n}{n_1, \ldots, n_k}$ способами.

Числа $\binom{n}{n_1, \ldots, n_k}$ являются коэффициентами в полиномиальной формуле

$$(x_1 + \ldots + x_k)^n = \sum_{n_1, \ldots, n_k} \binom{n}{n_1, \ldots, n_k} x_1^{n_1} \ldots x_k^{n_k}.$$

В частном случае $k=2$ верны равенства

$$\binom{n}{n_1, n_2} = \binom{n}{n_1} = \binom{n}{n_2} = \frac{n!}{n_1!n_2!},$$

и полиномиальная формула сводится к биномиальной формуле Ньютона.

К задаче о числе разбиений сводятся следующие задачи.

Задача 1. *Сколькими способами можно расселить 8 студентов по трем комнатам: одноместной, трехместной и четырехместной?*

Ответ: $\binom{8}{1, 3, 4} = \frac{8!}{1!\,3!\,4!} = 280$ способами.

Задача 2. *Сколькими способами четверо юношей могут пригласить танцевать четырех из шести девушек?*

Ответ: $\binom{6}{1, 1, 1, 1, 2} = \frac{6!}{1!\,1!\,1!\,1!\,2!} = 360$ способами.

Задача 3. *Сколькими способами можно раскрасить в красный, синий и зеленый цвет по 2 из 6 различных предметов?*

Ответ: $\binom{6}{2, 2, 2} = \frac{6!}{2!\,2!\,2!} = 90$ способами.

Задача 4. *Сколькими способами можно распределить 5 шаров 1, 2, 3, 4, 5 по трем ящикам 1, 2, 3 так, чтобы в них попало соответственно 1, 2, 3 шара?*

Ответ: $\binom{5}{1, 2, 2} = \frac{5!}{1!\,2!\,2!} = 30$ способами.

Задача 5. *Сколькими способами можно распределить 10 специалистов по четырем цехам 1, 2, 3, 4 так, чтобы в них попало соответственно 1, 2, 3, 4 специалиста?*

Ответ: $\binom{10}{1, 2, 3, 4} = \frac{10!}{1!\,2!\,3!\,4!} = 12\,600$ способами.

Рассмотрим n частиц, распределенных по k различным областям $1, 2, \ldots, k$ пространства.

Предположим, что частицы отмечены номерами $1, 2, \ldots, n$ и что в каждой области пространства их может находиться любое количество. В то же время будем считать, что состояние рассмат-

риваемой системы определяется числами $n_1, \ldots, n_k$ частиц в $1, \ldots, k$-й областях пространства, независимо от номеров частиц.

Условимся описанную модель называть *моделью Максвелла — Больцмана.*

Пример. Модель Максвелла — Больцмана для $n=3$ частиц и $k=2$ областей пространства описывается таблицей, в левой половине которой расположены номера частиц, находящихся в данной области, а в правой — число этих частиц.

Распределение		Состояние	
1	2	1	2
1 2 3		3	0
1 2	3	2	1
1 3	2		
2 3	1		
1	2 3	1	2
2	1 3		
3	1 2		
	1 2 3	0	3

Существует $k^n=2^3=8$ различных распределений $n=3$ частиц по $k=2$ областям. В состояние n_1, n_2 систему приводят $\binom{n}{n_1, n_2}$ распределений.

Задача 6. *Сколько распределений n частиц приводят систему в состояние $n_1, \ldots, n_k$ в модели Максвелла — Больцмана?*

Решение. Каждое такое распределение описывается $n_1, \ldots, n_k$-разбиением множества $A=\{1, \ldots, n\}$ номеров частиц. Поэтому число этих распределений равно числу $\binom{n}{n_1, \ldots, n^k}$ таких разбиений.

Задача 7. *Сколько всего существует различных распределений n частиц по k областям пространства в модели Максвелла — Больцмана?*

Решение. Используя полиномиальную формулу при $x_1=x_2= =\ldots=x_k=1$, убеждаемся в том, что число всех рассматриваемых распределений выражается равенством

$$\sum_{n_1, \ldots, n_k} \binom{n}{n_1 \ldots, n_k} = (1 + \ldots + 1)^n = k^n.$$

Задача 8. *Сколько всего существует возможных состояний системы в модели Максвелла — Больцмана?*

Ответ: $\binom{k+n-1}{n}$ состояний; в частности, если $n=3$ и $k=2$, то

$$\binom{2+3-1}{3}=\frac{4!}{3!\,1!}=4$$

в соответствии с примером.

По поводу решения задачи 8 см. решение задачи 6 из § 8.

§ 7. ПЕРЕСТАНОВКИ С ПОВТОРЕНИЯМИ

Каждое отображение f множества $A=\{1, 2, \ldots, n\}$ в множество $B=\{b_1, b_2, \ldots, b_k\}$, при котором прообразы $f^{-1}(b_1), \ldots, f^{-1}(b_k)$ элементов $b_1, \ldots, b_k$ состоят соответственно из $n_1, \ldots, n_k$ номеров, называется $n_1, \ldots, n_k$-перестановкой множества B $(n_1+\ldots+n_k=n)$.

Каждая $n_1, \ldots, n_k$-перестановка f множества $B=\{b_1, \ldots, b_k\}$ эквивалентна $n_1, \ldots, n_k$-разбиению $f^{-1}(b_1), \ldots, f^{-1}(b_k)$ множества $A=\{1, \ldots, n\}$. Поэтому число $n_1, \ldots, n_k$-перестановок множества B равно числу $\binom{n}{n_1, \ldots, n_k}$ всех $n_1, \ldots, n_k$-разбиений множества A.

Это предложение позволяет решить задачу о числе перестановок с повторениями.

Задача. *Сколькими способами можно переставить n предметов k различных типов $1, 2, \ldots, k$ по $n_1, \ldots, n_k$ одинаковых предметов, расположенных на n различных местах $1, 2, \ldots, n$?*

Ответ: $\binom{n}{n_1, \ldots, n_k}$ способами.

К задаче о числе перестановок с повторениями сводятся следующие задачи.

Задача 1. *Сколькими способами можно расположить в ряд 5 книг: 2 одинаковых романа и 3 одинаковых томика стихов?*

Ответ: $\binom{5}{2,\ 3}=\frac{5!}{2!\,3!}=10$ способами.

Задача 2. *Сколько сигналов можно составить, меняя порядок 7 флагов: 1 красный, 2 синих, 3 зеленых и 1 белый?*

Ответ: $\binom{7}{1,\ 2,\ 3,\ 1}=\frac{7!}{1!\,2!\,3!\,1!}=420$ сигналов.

Задача 3. *Сколько различных слов можно образовать, переставляя буквы в слове «анна»?*

Решение. $\binom{4}{2,\ 2}=\frac{4!}{2!\,2!}=6$ слов.

Вот эти слова: анна, ннаа, анан, аанн, наан, нана.

Задача 4. *Сколько слов можно составить из 5 букв a и не больше, чем из 3 букв b?*

Решение. Используя равенство задачи 9 из § 5, получаем:

$$\binom{5+0}{5,\ \ 0}+\binom{5+1}{5,\ \ 1}+\binom{5+2}{5,\ \ 2}+\binom{5+3}{5,\ \ 3}=\binom{5+0}{0}+$$

$$+\binom{5+1}{1}+\binom{5+2}{2}\binom{5+3}{3}=\binom{5+3+1}{3}=\frac{9!}{3!\,6!}=84.$$

Задача 5. *Для решения задачи с помощью вычислительной машины используются в определенном порядке по 2 программы каждого из 3 типов а, б, с. В списке указаны 88 последовательностей из 6 программ: ааббсс, аабсбс, аасббс, ...*

Все ли такие последовательности вошли в список?

Решение. Число всех рассматриваемых последовательностей программ выражается равенствами

$$\binom{6}{2,\ 2,\ 2}=\frac{6!}{2!\,2!\,2!}=90.$$

В имеющемся списке нет двух последовательностей.

§ 8. ВЫБОРКИ С ПОВТОРЕНИЯМИ

Каждое отображение f множества A из n элементов в множество N натуральных чисел, сумма $\sum\limits_{x\in A} f(x)$ значений $f(x)$ которого равна m, называется **выборкой (или сочетанием) с повторениями** m из n элементов множества A.

Число всех таких выборок выражается равенствами

$$\binom{n+m-1}{n-1}=\binom{n+m-1}{m}=\frac{(n+m-1)}{m!\,(n-1)!}.$$

Это предложение позволяет решить задачу о числе выборок с повторениями.

Задача. *Имеется по m одинаковых предметов каждого из n различных типов. Сколькими способами можно выбрать m из этих $n\cdot m$ предметов?*

Ответ: $\binom{n+m-1}{m}$ способами.

К задаче о числе выборок с повторениями сводятся следующие задачи.

Задача 1. *Сколькими способами можно выбрать 3 книги из 3 одинаковых романов и 3 одинаковых томиков стихов?*

Ответ: $\binom{2+3-1}{3}=4$ способами; среди выбранных могут быть 0, 1, 2, 3 экземпляра романа.

Задача 2. *Сколькими способами можно выбрать 13 из 52 стандартных карт, различая их только по масти?*

Ответ: $\binom{4+13-1}{13} = 560$ способами; способов намного меньше, чем при различии по масти и по значению.

Задача 3. *Сколько всего существует результатов опыта, заключающегося в подбрасывании 2 одинаковых игральных костей?*

Ответ: $\binom{6+2-1}{2} = 21$ результат; если кости различные, то 36.

Задача 4. *Сколькими способами можно выбрать 2 элемента из n различных пар одинаковых элементов?*

Ответ: $\binom{n+2-1}{2} = \binom{n+1}{2} = \frac{(n+1)\,n}{2}$ способами; n способами выбираются 2 одинаковые элемента, а $\binom{n}{2}$ способами — 2 разные элемента; по правилу Паскаля

$$n + \binom{n}{2} = \binom{n}{1} + \binom{n}{2} = \binom{n+1}{2}.$$

Замечание. Задача 4 имеет отношение к биологии: n аллельных генов $a_1, \ldots, a_n$ определяют $\binom{n+1}{2}$ генотипов, среди них n *гомозигот* $a_1a_1, \ldots, a_na_n$ и $\binom{n}{2}$ *гетерозигот* $a_1a_2, \ldots, a_{n-1}a_n$.

Задача 5. *Сколькими способами можно выбрать 3 из 12 букв А, А, А, Т, Т, Т, Г, Г, Г, Ц, Ц, Ц?*

Ответ: $\binom{4+3-1}{3} = 20$ способами.

Замечание. Буквы А, Т, Г, Ц обозначают нуклеотиды Аденин, Тимин, Гуанин, Цитозин. Число троек нуклеотидов оказывается равным магическому числу 20 — числу стандартных аминокислот, на которые разлагаются молекулы белка. На этом равенстве основывался простейший белковый код, предложенный физиком Гамовым.

Рассмотрим n частиц, распределенных по k различным областям $1, \ldots, k$ пространства.

Предположим, что частицы одинаковы и что в каждой области пространства их может находиться любое количество. Состояние рассматриваемой системы определяется числами $n_1, \ldots, n_k$ частиц в $1, \ldots, k$-й областях пространства.

Условимся описанную модель называть моделью Бозе — Эйнштейна. Она отличается от модели Максвелла — Больцмана предположением об одинаковости частиц. В модели Бозе — Эйнштейна не существует различных распределений частиц, определяющих одинаковые состояния системы. А в модели Максвелла — Больцмана с различными частицами такие распределения существуют.

Задача 6. *Сколько всего существует возможных состояний системы в модели Бозе — Эйнштейна?*

Решение. Число состояний системы в модели Бозе — Эйнштейна, как и в модели Максвелла — Больцмана, равно числу всех семейств k натуральных чисел $n_1, \ldots, n_k$, в сумме равных n:

$$n_1+n_2+\ldots+n_k=n.$$

Каждое такое семейство является выборкой с повторениями n из k элементов множества $A=\{1, \ldots, k\}$. Следовательно, число состояний системы в модели Бозе — Эйнштейна, как и в модели Максвелла — Больцмана, равно $\binom{k+n-1}{n}$.

Глава 2*

ЗАДАЧА О СУММЕ СТЕПЕНЕЙ

Задача о сумме степеней последовательных натуральных чисел интересна в различных отношениях. При ее решении используются комбинаторные приемы и метод разностных уравнений. Они могут применяться также для решения многих других задач.

При первом чтении эту главу можно пропустить.

§ 1. МАТЕМАТИЧЕСКИЙ АППАРАТ

Прежде, чем сформулировать задачу о сумме степеней и решить ее, подготовим дополнительный математический аппарат.

1.1. Разностное уравнение $\Delta g=f$

В некоторых случаях последовательность $\Delta g=f$, значениями которой являются разности $\Delta g(n)=g(n+1)-g(n)$ значений числовой последовательности g, проще поддается исследованию, чем сама последовательность g. Поэтому важно уметь по последовательности $\Delta g=f$ разностей определять последовательность g, т. е. решать разностное уравнение

$$\Delta g=f.$$

Пример 1. Рассмотрим арифметическую прогрессию g с начальным членом a и шагом b. Для каждого номера $n=0, 1, 2, \ldots$ значение $g(n)$ последовательности g определяется равенством $g(n)=a+bn$. Последовательность Δg в этом случае является постоянной последовательностью f, каждое значение которой равно b. В самом деле, для каждого номера n

$$\Delta g(n)=g(n+1)-g(n)=[a+b(n+1)]-[a+bn]=b.$$

Другими словами, арифметическая прогрессия с начальным членом a и шагом b является решением разностного уравнения

$$\Delta g=f \quad (f(n)=b,\ n=0, 1, 2, \ldots). \tag{1'}$$

Заметим, что

$$g(0)=a. \qquad (2')$$

В связи с этим говорят, что арифметическая прогрессия с начальным членом a и шагом b является решением разностного уравнения (1′), удовлетворяющим начальному условию (2′).

Пример 2. Рассмотрим произвольное число c и натуральное число m. Определим число $c^{(m)}$ равенствами:

$$c^{(m)}=0 \quad (m<0);$$

$$c^{(0)}=1,\ c^{(1)}=c, \quad c^{(2)}=c(c-1);$$

$$c^{(m)}=c(c-1)\cdot\ldots\cdot(c-m+1) \quad (m=3,\ 4,\ 5,\ \ldots).$$

Число $c^{(m)}$ связано с биномиальным коэффициентом $\binom{c}{m}$ равенством

$$\binom{c}{m}=\frac{c^{(m)}}{m!}.$$

Рассмотрим последовательность f со значениями $f(n)=n^{m}$ $(n=0,\ 1,\ 2,\ \ldots)$ и последовательность s со значениями

$$s(0)=0,$$

$$s(n)=\sum_{k=0}^{n-1} f(k)=\sum_{k=0}^{n-1} k^{(m)} \quad (n=1, 2, 3, \ldots).$$

Нетрудно проверить, что

$$\Delta s(0)=s(1)-s(0)=f(0)=0^{(m)},$$

$$\Delta s(n)=s(n+1)-s(n)=f(n)=n^{(m)} \quad (n=1,\ 2,\ 3,\ldots).$$

Таким образом, последовательность $g=s$ является решением разностного уравнения

$$\Delta g=f \quad (f(n)=n^{(m)},\ n=0,\ 1,\ 2,\ \ldots), \qquad (1'')$$

удовлетворяющим начальному условию

$$g(0)=0. \qquad (2'')$$

Рассмотрим также последовательность p со значениями

$$p(n)=\frac{1}{m+1}\,n^{(m+1)} \quad (n=0, 1, 2, \ldots).$$

В частности,

$$p(0)=\frac{1}{m+1}\,0^{(m+1)}=0.$$

Для $m=2,\ 3,\ 4,\ldots$ верны равенства

$$\Delta p(n)=p(n+1)-p(n)=\frac{1}{m+1}(n+1)^{(m+1)}=$$

$$=\frac{1}{m+1}\left[(n+1)\,n\ldots(n+1-m)-n(n-1)\ldots(n-m)\right]=$$

$$=\frac{1}{m+1}\,n^{(m)}\left[(n+1)-(n-m)\right]=n^{(m)} \quad (n=0, 1, 2, \ldots).$$

Нетрудно проверить (упражнение), что равенства $\Delta p(n) = n^{(m)}$ $(n=0, 1, 2,...)$ верны и для $m=0, 1$. Сказанное означает, что последовательность $g=p$ является решением разностного уравнения (1''), удовлетворяющим начальному условию (2'').

Таким образом, каждая из последовательностей s и p является решением разностного уравнения (1''), удовлетворяющим начальному значению (2''). Естественно, возникает вопрос: не следует ли отсюда, что последовательности s и p равны? Из теоремы о существовании и единственности решения для разностного уравнения $\Delta g = f$, которая будет доказана в этом пункте, вытекает, что это действительно так. Поэтому верны равенства:

$$\sum_{k=0}^{n-1} k^{(m)} = \frac{1}{m+1} n^{(m+1)} \quad (n = 1, 2, 3, \ldots).$$

Получено удобное выражение для вычисления сумм чисел $k^{(m)}$. Примеры 1, 2 подводят к следующим общим определениям. Рассмотрим произвольные числовые последовательности f, g и число c.

Обозначим символом Δg последовательность со значениями

$\Delta g(n) = g(n+1) - g(n)$ $(n=0, 1, 2,...)$.

Если верно равенство

$$\Delta g = f, \tag{1}$$

то будем говорить, что последовательность g является «решением разностного уравнения (1)». Если, кроме того, верно равенство

$$g(0) = c, \tag{2}$$

то условимся говорить, что последовательность g является «решением разностного уравнения (1), удовлетворяющим начальному условию (2)».

Последовательность f называется «правой частью разностного уравнения (1)».

Замечание. Равенство $\Delta g = f$ для последовательностей верно тогда и только тогда, когда для каждого номера n верно равенство $\Delta g(n) = f(n)$ для чисел:

$$\Delta g(n) = f(n) \quad (n=0, 1, 2,...).$$

Используя принцип индукции, нетрудно доказать, что верна следующая теорема о существовании и единственности решения для разностного уравнения (1).

Теорема. *Последовательность $g=s$ со значениями*

$$s(0) = c, \quad s(n) = c + \sum_{k=0}^{n-1} f(k) \quad (n = 1, 2, \ldots) \tag{3}$$

является единственным решением разностного уравнения (1), *удовлетворяющим условию* (2).

Доказательство. а. Последовательность $g=s$, значения которой определены равенствами (3), является решением разностного урав-

нения (1), удовлетворяющим начальному условию (2). В самом деле,

$$\Delta s(0)=s(1)-s(0)=[c+f(0)]-c=f(0),$$

$$\Delta s(n)=s(n+1)-s(n)=\left[c+\sum_{k=0}^{n} f(k)\right]-$$

$$-\left[c+\sum_{k=0}^{n-1} f(k)\right]=f(n) \quad (n=1, 2, 3, \ldots).$$

Кроме того, по определению,

$$s(0)=c.$$

б. Рассмотрим произвольное решение $g=p$ уравнения (1), удовлетворяющее начальному условию (2). Используя принцип индукции и результат пункта 1 доказательства, покажем, что $p=s$.

Рассмотрим множество M всех номеров m, для которых верно равенство

$$p(m)=s(m).$$

1. Номер 0 принадлежит множеству M, так как

$$p(0)=c=s(0).$$

2. Для каждого номера n верно предложение: если n принадлежит M, то $n+1$ принадлежит M. В самом деле, если номер n принадлежит множеству M, то

$$p(n)=s(n).$$

В то же время, так как p и s являются решениями разностного уравнения (1),

$$p(n+1)-p(n)=s(n+1)-s(n).$$

Отсюда вытекает, что

$$p(n+1)-s(n+1)=p(n)-s(n)=0.$$

И, следовательно,

$$p(n+1)=s(n+1),$$

т. е. номер $n+1$ принадлежит множеству M.

Таким образом, условия (1) и (2) принципа индукции выполнены, и множество M совпадает с множеством N всех натуральных чисел. Это значит, что равенство $p(n)=s(n)$ верно для каждого номера n, т. е. $p=s$.

Теорема доказана.

Как было отмечено при рассмотрении примера 2, из доказанной теоремы вытекает

Следствие. Для каждых натуральных чисел $m\geqslant 0$ и $n>0$ верно равенство

$$\sum_{k=0}^{n-1} k^{(m)}=\frac{1}{m+1}n^{(m+1)}.$$

Доказательство. Рассмотрим последовательности f, s и p со значениями соответственно

$$f(n) = n^{(m)} \quad (n = 0, 1, 2, \ldots),$$

$$s(0) = 0, \quad s(n) = \sum_{k=0}^{n-1} k^{(m)} \quad (n = 1, 2, \ldots),$$

$$p(n) = \frac{1}{m+1} n^{(m+1)} \quad (n = 0, 1, 2, \ldots).$$

Как было показано, каждая из последовательностей $g = s$ и $g = p$ является решением разностного уравнения (1″), удовлетворяющим условию 2 при $c = 0$. По теореме о существовании и единственности решения для разностного уравнения (1″) отсюда следует, что $p = s$.

2.1. Треугольная система линейных уравнений

Рассмотрим систему линейных уравнений специального треугольного вида, благодаря которому эта система легко решается.

Пример. Рассмотрим семейства чисел:

$$\begin{pmatrix} a_{11} & & \\ a_{21} & a_{22} & \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

Если верны равенства

$$\left.\begin{aligned} &a_{11}x_1 = b_1, \\ &a_{21}x_1 + a_{22}x_2 = b_2, \\ &a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3, \end{aligned}\right\} \tag{4′}$$

то говорят, что семейство чисел x_1, x_2, x_3 является решением системы линейных уравнений (4′). Предположим, что

$$a_{11} \neq 0, \; a_{22} \neq 0, \; a_{33} \neq 0. \tag{5′}$$

В этом случае, используя треугольный вид системы (4′) и последовательно решая составляющие ее уравнения, можно легко найти единственное решение x_1, x_2, x_3 этой системы.

В самом деле, если верны равенства (4′), то

$$x_1 = \frac{1}{a_{11}} b_1. \tag{6′}$$

Это равенство вместе с равенством

$$x_2 = \frac{1}{a_{22}} [b_2 - a_{21}x_1] \tag{6′}$$

определяют x_2. Равенства для x_1 и x_2 вместе с равенством

$$x_3 = \frac{1}{a_{33}} [b_3 - (a_{31}x_1 - a_{32}x_2)] \tag{6′}$$

определяют x_3. Таким образом, если решение x_1, x_2, x_3 системы (4′) существует, то оно единственно и определяется равенствами (6′).

Подставляя определенные равенствами (6′) числа x_1, x_2, x_3 в равенства (4′), убеждаемся в том, что эти числа действительно образуют решение рассматриваемой системы линейных уравнений:

$$a_{11}\left(\frac{1}{a_{11}}b_1\right)=b_1,$$

$$a_{21}x_1+a_{22}\left(\frac{1}{a_{22}}[b_2-a_{21}x_1]\right)=b_2,$$

$$a_{31}x_1+a_{32}x_2+a_{33}\left(\frac{1}{a_{33}}[b_3-(a_{31}x_1+a_{32}x_2)\right)=b_3.$$

В общем случае картина аналогична.

Рассмотрим произвольное натуральное число $m>0$ и семейства чисел:

$$\begin{Bmatrix} a_{11} \\ a_{21}a_{22} \\ \cdots\cdots\cdots \\ a_{i1}a_{i2}\ldots a_{ij}\ldots a_{ii} \\ \cdots\cdots\cdots \\ a_{m1}a_{m2}\ldots a_{mj}\ldots a_{mm} \end{Bmatrix}, \quad \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_m \end{Bmatrix}, \quad \begin{Bmatrix} b_1 \\ b_2 \\ \vdots \\ b_i \\ \vdots \\ b_m \end{Bmatrix}.$$

В сокращенной записи эти семейства обозначаются соответственно символами

$$(a_{ij})_{1\leqslant i\leqslant m,\ 1\leqslant j\leqslant i}, \quad (x_i)_{1\leqslant i\leqslant m}, \quad (b_i)_{1\leqslant i\leqslant m}.$$

Вместо $(x_i)_{1\leqslant i\leqslant m}$ будем писать также $x_1, \ldots, x_m$. Если верны равенства

$$\left.\begin{array}{l} a_{11}x_1=b_1, \\ a_{21}x_1+a_{22}x_2=b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{i1}x_1+a_{i2}x_2+\ldots+a_{ij}x_j+\ldots+a_{ii}x_i=b_i, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mj}x_j+\ldots+a_{mm}x_m=b_m, \end{array}\right\} \qquad (4)$$

имеющие в сокращенной записи вид

$$\sum_{j=1}^{i} a_{ij}x_j=b_i \quad (i=1,\ldots,m), \qquad (4a)$$

то говорят, что «семейство чисел $(x_i)_{1\leqslant i\leqslant m}$ является решением системы линейных уравнений с семейством коэффициентов $(a_{ij})_{1\leqslant i\leqslant m,\ 1\leqslant j\leqslant i}$ и семейством правых частей $(b_i)_{1\leqslant i\leqslant m}$». Коротко то же самое выражают фразой «семейство чисел $x_1, \ldots, x_m$ является решением системы линейных уравнений (4)» или фразой «семейство $x_1, \ldots, x_m$ является решением системы (4)».

Предположим, что

$$a_{ii} \neq 0 \quad (i=1, \ldots, m). \tag{5}$$

В этом случае, используя треугольный вид системы (4) и последовательно решая составляющие ее уравнения, можно легко найти ее единственное решение $x_1, \ldots, x_m$. Формально это делается следующим образом. Чтобы сократить запись, условимся считать, что

$$\sum_{j=1}^{0} a_{1j} x_j = 0.$$

Если $m > 1$ и $k = 1, \ldots, m-1$, то рассмотрим произвольные числа $x_1, \ldots, x_k$. Определим число x_{k+1} равенством

$$x_{k+1} = \frac{1}{a_{k+1,k+1}} \left[b_{k+1} - \sum_{j=1}^{k} a_{k+1,j} x_j \right] \quad (k = 0, \ldots, m-1). \tag{6a}$$

В развернутой форме равенства (6) имеют вид:

$$\left.\begin{array}{l}
x_1 = \dfrac{1}{a_{11}} b_1, \\
x_2 = \dfrac{1}{a_{22}} [b_2 - a_{21} x_1], \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
x_{k+1} = \dfrac{1}{a_{k+1,k+1}} [b_{k+1} - (a_{k+1,1} x_1 + \ldots + a_{k+1,j} x_j + \ldots + a_{k+1,k} x_k)], \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
x_m = \dfrac{1}{a_{mm}} [b_m - (a_{m1} x_1 + \ldots + a_{mj} x_j + \ldots + a_{m,m-1} x_{m-1})].
\end{array}\right\} \tag{6}$$

Теорема. *Если выполнены условия* (5), *то равенства* (6) *определяют семейство чисел* $x_1, \ldots, x_m$, *которое является единственным решением системы линейных уравнений* (4).

Доказательство. Используем принцип индукции. Предположим, что условия (5) выполнены, и рассмотрим множество M всех чисел $n = m-1$, для которых равенства (6) определяют семейство чисел $x_1, \ldots, x_m$, являющееся единственным решением системы уравнений (4).

1. Ясно, что $0 \in M$: равенство

$$x_1 = \frac{1}{a_{11}} b_1$$

определяет число x_1, которое является единственным решением линейного уравнения

$$a_{11} x_1 = b_1.$$

2. Докажем, что для каждого натурального числа n верно предложение: если $n \in M$, то $n+1 \in M$.

Предположим, что при $m=n$ равенства (6) определяют семейство чисел $x_1, \ldots, x_n$, которое является единственным решением системы линейных уравнений (4).

Рассмотрим число x_{n+1}, определяемое равенством (6) при $k=n$. Семейство чисел $x_1, \ldots, x_n, x_{n+1}$ является единственным решением системы линейных уравнений (4) при $m=n+1$. В самом деле, по предположению, семейство чисел $x_1, \ldots, x_n$ является решением системы (4) при $m=n$. Кроме того,

$$\sum_{j=1}^{n} a_{n+1,j}x_j + a_{n+1,n+1}x_{n+1} = \sum_{j=1}^{n} a_{n+1,j}x_j +$$

$$+ a_{n+1,n+1}\frac{1}{a_{n+1,n+1}}\left[b_{n+1} - \sum_{j=1}^{n} a_{n+1,j}x_j\right] = b_{n+1}.$$

Следовательно, семейство $x_1, \ldots, x_n, x_{n-1}$ является решением системы (4) при $m=n+1$.

Рассмотрим произвольное семейство чисел $y_1, \ldots, y_n, y_{n+1}$, для которого верны равенства

$$\sum_{j=1}^{i} a_{ij}y_j = b_i \quad (i = 1, \ldots, n+1).$$

Из сделанного относительно $x_1, \ldots, x_n$ предположения вытекает, что $y_1=x_1, \ldots, y_n=x_n$. Следовательно,

$$y_{n+1} = \frac{1}{a_{n+1,n+1}}\left[b_{n+1} - \sum_{j=1}^{n} a_{n+1,j}y_j\right] =$$

$$= \frac{1}{a_{n+1,n+1}}\left[b_{n+1} - \sum_{j=1}^{n} a_{n+1,j}x_j\right] = x_{n+1}.$$

Таким образом, семейство $x_1, \ldots, x_n, x_{n+1}$ является единственным решением системы (4) при $m=n+1$.

Условия (1) и (2) принципа индукции выполнены. Значит, $M=N$. Теорема доказана.

§ 2. ПОСТАНОВКА ЗАДАЧИ

Сначала сформулируем задачу о сумме степеней описательно, затем уточним формулировку задачи и сведем дело к решению некоторого разностного уравнения.

1.2. Примеры

Прежде, чем описать задачу о сумме степеней в общем виде, рассмотрим несколько простых примеров.

Пример 1. По формуле суммы арифметической прогрессии для каждого натурального числа n верны равенства

$$\sum_{0\leqslant k<n} k = \frac{(n-1)\,n}{2} = \frac{1}{2}n^2 - \frac{1}{2}n.$$

В частности, если $n=0$, то эти равенства эквивалентны равенству $0=0$.

Пример 2. Используя принцип математической индукции, нетрудно доказать, что для каждого натурального числа n верны равенства

$$\sum_{0\leqslant k<n} k^2 = \frac{(n-1)\,n\,(2n-1)}{6} = \frac{1}{3}n^3 - \frac{1}{2}n^2 + \frac{1}{6}n.$$

Рассмотрим множество M всех натуральных чисел m, для которых верно равенство

$$\sum_{0\leqslant k<m} k^2 = \frac{(m-1)\,m\,(2m-1)}{6}.$$

1. Число $0\in M$, так как при $m=0$ это равенство эквивалентно равенству $0=0$.

2. Для каждого натурального числа n, если $n\in M$, $n+1\in M$. В самом деле, для каждого натурального числа n верно равенство

$$\sum_{0\leqslant k<n+1} k^2 = \sum_{0\leqslant k<n} k^2 + n^2.$$

Если $n\in M$, то

$$\sum_{0\leqslant k<n+1} k^2 = \frac{(n-1)\,n\,(2n-1)}{6} + n^2 = \frac{((n+1)-1)\,(n+1)\,(2\,(n+1)-1)}{6},$$

и значит, $n+1\in M$.

Таким образом, условия принципа индукции выполнены и множество M совпадает с множеством N всех натуральных чисел. Рассматриваемые равенства верны для каждого натурального числа n.

Пример 3. Используя принцип математической индукции, нетрудно доказать, что для каждого натурального числа n верны равенства

$$\sum_{0\leqslant k<n} k^3 = \left(\frac{(n-1)\,n}{2}\right)^2 = \frac{1}{4}n^4 - \frac{1}{2}n^3 + \frac{1}{4}n^2.$$

Рассмотрим множество M всех натуральных чисел m, для которых верно равенство

$$\sum_{0 \leqslant k < m} k^3 = \left(\frac{(m-1)\,m}{2}\right)^2.$$

1. Число $0 \in M$, так как при $m=0$ это равенство эквивалентно равенству $0=0$.

2. Для каждого натурального числа n, если $n \in M$, $n+1 \in M$. В самом деле, для каждого натурального числа n верно равенство

$$\sum_{0 \leqslant k < n+1} k^3 = \sum_{0 \leqslant k < n} k^3 + n^3.$$

Если $n \in M$, то

$$\sum_{0 \leqslant k < n+1} k^3 = \left(\frac{(n-1)\,n}{2}\right)^2 + n^3 = \left(\frac{((n+1)-1)\,(n+1)}{2}\right)^2,$$

и значит, $n+1 \in M$.

Таким образом, условия принципа индукции выполнены и множество M совпадает с множеством N всех натуральных чисел. Рассматриваемые равенства верны для каждого натурального числа n.

Замечание. Из равенства примеров 1 и 3 следует неожиданное, на первый взгляд, равенство

$$\sum_{0 \leqslant k < n} k^3 = \left(\sum_{0 \leqslant k < n} k\right)^2.$$

(По поводу этого равенства см. книгу Д. Пойа «Как решить задачу». М., Учпедгиз, 1961.)

Пример 4. Рассмотрим произвольные числа a и b, арифметическую прогрессию f с начальным членом a и шагом b

$$f(n) = a + bn \quad (n = 0, 1, 2, \ldots)$$

и последовательность s со значениями

$$s(0) = 0,$$

$$s(n) = \sum_{k=0}^{n-1} f(k) \quad (n = 1, 2, 3, \ldots).$$

Вычислять значения $s(n)$ последовательным сложением чисел $f(k) = a + bk$ трудно для больших номеров n. Поэтому возникает задача: найти выражение для суммы $s(n)$, которое позволяло бы сравнительно просто вычислять эту сумму.

Используя результат примера 1, находим:

$$s(n)=\sum_{k=0}^{n-1}(a+bk)=na+b\sum_{k=0}^{n-1}k=na+b(n-1)n/2$$
$$(n=1,2,3,\ldots).$$

Пример 5. Рассмотрим произвольные числа a, b и c, последовательность f со значениями

$$f(n)=a+bn+cn^2 \quad (n=0,\ 1,\ 2,\ \ldots)$$

и последовательность s со значениями

$$s(0)=0,$$

$$s(n)=\sum_{k=0}^{n-1}f(k) \quad (n=1,2,3,\ldots).$$

Вычислять значения $s(n)$ последовательным сложением чисел $f(k)=a+bk+ck^2$ еще труднее для бо́льших номеров n, чем в примере 4. Поэтому в данном случае задача о простом выражении для суммы $s(n)$ еще важнее.

Используя результаты примеров 2 и 4, находим:

$$s(n)=\sum_{k=0}^{n-1}(a+bk+ck^2)=\sum_{k=0}^{n-1}(a+bk)+c\sum_{k=0}^{n-1}k^2=$$
$$=na+b(n-1)n/2+c(n-1)n(2n-1)/6 \quad (n=1,2,3,\ldots).$$

Примеры 1—5 подводят к следующей общей задаче. Рассмотрим натуральное число $m>0$, последовательность f со значениями

$$f(n)=n^m \quad (n=0,\ 1,\ 2,\ \ldots) \tag{1}$$

и последовательность s со значениями

$$s(0)=0, \tag{2}$$

$$s(n)=\sum_{k=0}^{n-1}f(k)=\sum_{k=0}^{n-1}k^m \quad (n=1,2,3,\ldots). \tag{3}$$

Задача: найти выражение для суммы $s(n)$, которое позволяло бы сравнительно просто вычислять эту сумму.

Так как неясно, что значит «сравнительно просто вычислять», то данное описание задачи нуждается в уточнении.

2.2. Формулировка задачи

В примерах 1, 2, 3 пункта 1.2 при $m=1$, 2, 3 для суммы $s(n)$ можно получить выражения в виде значений полиномов ($n=1$, 2, 3, ...):

$$s(n)=\sum_{k=0}^{n-1} k^1=\frac{1}{2}n^2-\frac{1}{2}n \qquad (m=1),$$

$$s(n)=\sum_{k=0}^{n-1} k^2=\frac{1}{3}n^3-\frac{1}{2}n^2+\frac{1}{6}n \qquad (m=2),$$

$$s(n)=\sum_{k=0}^{n-1} k^3=\frac{1}{4}n^4-\frac{1}{2}n^3+\frac{1}{4}n^2 \qquad (m=3).$$

Эти равенства наводят на мысль о том, что для каждого натурального числа $m>0$ существует полином p степени $m+1$, для которого верны равенства

$$p(n)=s(n) \quad (n=0,\ 1,\ 2,\ \ldots). \tag{4}$$

Значения полинома p вычислять сравнительно проще, чем значения суммы $s(n)$ последовательным сложением составляющих ее чисел. Поэтому задачу можно сформулировать следующим образом.

Задача о сумме степеней. *Для каждого натурального числа $m>0$ найти полином p степени $m+1$, для которого верны равенства*

$$p(0)=0,$$

и

$$p(n)=\sum_{k=0}^{n-1} k^m \qquad (n=1, 2, 3, \ldots).$$

Полином p степени $m+1$ определяется своими коэффициентами a_0, a_1, ..., a_{m+1}. Задача о сумме степеней, следовательно, сводится к задаче: найти числа a_0, a_1,..., a_{m+1}, для которых верны равенства $a_0=0$ и

$$\sum_{k=0}^{n-1} k^m=a_0+a_1n+\ldots+a_{m+1}n^{m+1} \quad (n=1, 2, 3, \ldots). \tag{5}$$

Для случаев $m=1$, 2, 3 числа a_0, a_1, ..., a_{m+1} были найдены:

$$a_0=0, \quad a_1=-\frac{1}{2}, \quad a_2=\frac{1}{2} \qquad (m=1);$$

$$a_0=0, \quad a_1=\frac{1}{6}, \quad a_2=-\frac{1}{2}, \quad a_3=\frac{1}{3} \qquad (m=2);$$

$$a_0=0, \quad a_1=0, \quad a_2=\frac{1}{4}, \quad a_3=-\frac{1}{2}, \quad a_4=\frac{1}{4} \quad (m=3).$$

Равенство $a_0=0$ эквивалентно условию $p(0)=0$. Для старших коэффициентов верно равенство

$$a_{m+1}=\frac{1}{m+1}.$$

Можно предположить, что это равенство верно для каждого натурального числа $m>0$.

3.2. Разностное уравнение

Рассуждая как в примере 2 пункта 1.1, убеждаемся, что последовательность $g=s$, значения которой определяются равенствами (2) и (3), является решением разностного уравнения

$$\Delta g=f \quad (f(n)=n^m,\ n=0,\ 1,\ 2, \ldots), \tag{1''}$$

удовлетворяющим начальному условию

$$g(0)=0. \tag{2''}$$

Из теоремы о существовании и единственности решения для таких уравнений, доказанной в пункте 1.1, следует, что задача о сумме степеней сводится к поиску «имеющего форму полинома» решения разностного уравнения (1″). Точнее, задача о сумме степеней эквивалентна следующей задаче: *для каждого натурального числа* $m>0$ *найти полином* p *степени* $m+1$, *для которого последовательность* g *со значениями*

$$g(n)=p(n) \quad (n=0,\ 1,\ 2, \ldots) \tag{6}$$

является решением разностного уравнения (1″), *удовлетворяющим начальному условию* (2″).

Эту последнюю задачу и будем решать.

§ 3. РЕШЕНИЕ ЗАДАЧИ

Как и задача о сумме степеней, эквивалентная ей задача об имеющем форму полинома решении уравнения (1″) сводится к определению коэффициентов $a_1, \ldots, a_m$ полинома p, значения которого удовлетворяют равенствам (6). Для определения этих коэффициентов используется способ, называемый «методом неопределенных коэффициентов». Поступаем следующим образом: 1) используя равенство (1″), составляем систему линейных уравнений, решением которой являются искомые коэффициенты $a_1, \ldots, a_m$ полинома p; 2) решаем полученную систему линейных уравнений и определяем коэффициенты $a_1, \ldots, a_m$. При этом последовательность g, значения которой определяются равенствами (6), является решением разностного уравнения (1″), удовлетворяющим начальному условию (2″). Тем самым задача решена.

1.3. Составление системы линейных уравнений

Рассмотрим произвольный полином p степени $m+1$:

$$p(x)=\sum_{l=0}^{m+1} a_l x^l$$

со старшим коэффициентом

$$a_{m+1}=\frac{1}{m+1}$$

и свободным членом $a_0=0$. Коэффициенты $a_1, \ldots, a_m$ оставим пока неопределенными.

Если полином p является решением поставленной задачи (т. е. последовательность g со значениями, определяемыми равенствами (6), является решением разностного уравнения (1″)), то верны равенства

$$\Delta p(n)=n^m \quad (n=0,\ 1,\ 2, \ldots).$$

Числа $\Delta p(n)$ являются значениями некоторого полинома Δp степени m. Точнее, имеет место следующее предложение.

Лемма. *Верны равенства*

$$\Delta p(n) = \sum_{r=0}^{m} c_r n^r \quad (r = 0, 1, 2, \ldots),$$

где

$$c_r = \sum_{l=r+1}^{m+1} \binom{l}{r} a_l \quad (r = 0, 1, \ldots, m).$$

Доказательство. 1. Прежде всего заметим, что для каждого натурального числа n верно равенство

$$\Delta p(n) = \sum_{l=1}^{m+1} a_l \Delta(n^l), \qquad (*)$$

где

$$\Delta(n^l)=(n+1)^l-n^l \quad (l=1, \ldots, m+1).$$

В самом деле,

$$\Delta p(n) = p(n+1) - p(n) = \sum_{l=1}^{m+1} a_l (m+1)^l - \sum_{l=1}^{m+1} a_l n^l =$$

$$= \sum_{l=1}^{m+1} a_l\ [(n+1)^l - n^l] = \sum_{l=1}^{m+1} a_l \Delta(n^l).$$

2. Используя формулу Ньютона для бинома, получаем:

$$(n+1)^l = \sum_{k=0}^{l} \binom{l}{k} n^k = n^l + \sum_{k=0}^{l-1} \binom{l}{k} n^k \quad (l = 1, \ldots, m+1),$$

откуда

$$\Delta (n)^l = \sum_{k=0}^{l-1} \binom{l}{k} n^k \quad (l = 1, \ldots, m+1).$$

Отсюда и из равенства (*) следует, что

$$\Delta p(n) = \sum_{l+1}^{m+1} \left[a_l \sum_{k=0}^{l-1} \binom{l}{k} n^k\right] = \sum_{l=1}^{m+1} \left[\sum_{k=0}^{l-1} \binom{l}{k} a_l n^k\right]. \qquad (**)$$

3. Суммируя в правой части равенства (**) слагаемые, содержащие одинаковые степени n^r, получаем нужные равенства для коэффициентов c^r $(r=0, 1, \ldots, m)$. Заметим, что степень n^r

содержится в слагаемом $\binom{l}{k} a_l n^k$ тогда и только тогда, когда $l \geqslant r+1$ и $k=r$. Объединим все эти слагаемые в группу. Сумма их равна числу

$$\sum_{l=r+1}^{m+1} \binom{l}{r} a_l n^r.$$

При этом каждое слагаемое $\binom{l}{k} a_l n^k$ попадает в одну и только одну такую группу. Поэтому общая сумма получается сложением групповых сумм. Другими словами, верно равенство

$$\sum_{l=1}^{m+1} \left[\sum_{k=0}^{l-1} \binom{l}{k} a_l n^k \right] = \sum_{r=0}^{m} \left[\sum_{l=r+1}^{m+1} \binom{l}{r} a_l n^r \right]. \qquad (***)$$

Вместе с тем верны равенства

$$c_r n^r = n^r \sum_{l=r+1}^{m+1} \binom{l}{r} a_l = \sum_{l=r+1}^{m+1} \binom{l}{r} a_l n^r, \qquad (****)$$

где c_r — числа, фигурирующие в формулировке леммы.

Из равенств (*) — (****) следует равенство

$$\Delta p(n) = \sum_{r=0}^{m} c_r n^r.$$

Лемма доказана.

Из доказанной леммы следует, что равенство $\Delta p(n) = n^m$ эквивалентно равенству

$$\sum_{r=0}^{m} c_r n^r = n^m.$$

Это равенство верно, если верны равенства $c_m = 1$, $c_r = 0$ $(r = 0, \ldots$ $\ldots, m-1)$. Первое из этих равенств верно:

$$c_m = \binom{m+1}{m} a_{m+1} = (m+1) \frac{1}{m+1} = 1.$$

Таким образом, дело сводится к следующей задаче: определить коэффициенты $a_1, \ldots, a_m$ полинома так, чтобы были верны равенства

$$\sum_{l=r+1}^{m+1} \binom{l}{r} a_l = 0 \quad (r = 0, \ldots, m-1), \qquad (7)$$

т. е. мы свели задачу к решению некоторой системы линейных уравнений.

Замечание. В проведенном только что рассуждении используется очевидное утверждение: если верны равенства

$$c_m = 1, \quad c_r = 0 \;\; (r = 0, \ldots, m-1),$$

то верны равенства

$$\sum_{r=0}^{m} c_r n^r = n^m \quad (n = 0, 1, 2, \ldots).$$

Верно также менее очевидное обратное утверждение о том, что из вторых равенств следуют первые. Это вытекает, например, из следующего предложения о равенстве полиномов. Рассмотрим произвольные полиномы

$$g(x) = \sum_{r=0}^{m} c_r x^r, \qquad h(x) = \sum_{r=0}^{m} d_r x^r$$

степени $m > 0$. Если существуют $m+1$ различные числа x_n, для которых эти полиномы имеют равные значения:

$$g(x_n) = h(x_n) \quad (n = 0, 1, \ldots, m),$$

то равны соответствующие коэффициенты этих полиномов:

$$c_r = d_r \ (r = 0, 1, \ldots, m).$$

(Энциклопедия элементарной математики. Кн. II. Алгебра. М., Гостехиздат, 1951: ст. Л. Я. Окунева «Кольцо многочленов и рациональных функций», гл. 1, § 1, условие равенства многочленов; § 3, теорема 10, следствие).

В рассматриваемом случае

$$h(x) = x^m,$$
$$d_m = 1, \quad d_r = 0 \ (r = 0, \ldots, m-1),$$
$$x_n = n \ (n = 0, 1, \ldots, m).$$

2.3. Решение системы линейных уравнений (7)

Система линейных уравнений (7) легко сводится к эквивалентной треугольной системе линейных уравнений. Заметим, что при $l = m+1$ верно равенство

$$\binom{l}{r} a_l = \frac{1}{m+1}\binom{m+r}{r}.$$

Введем следующие обозначения:

$$\left.\begin{aligned} & i = m - r, \quad j = m + 1 - l; \\ & a_{ij} = \binom{m+1-j}{m-i} \quad (i = 1, \ldots, m; \ j = 1, \ldots, i); \\ & x_j = a_{m+1-j}; \\ & b_i = -\frac{1}{m+1}\binom{m+i}{m-1}. \end{aligned}\right\} \quad (8)$$

В этих новых обозначениях система (7) имеет вид

$$\sum_{j=1}^{i} a_{ij}x_j - b_i = 0 \qquad (i=1, \ldots, m).$$

Следовательно, она эквивалентна треугольной системе линейных уравнений

$$\sum_{j=1}^{i} a_{ij}x_j = b_i \qquad (i=1, \ldots, m), \tag{9}$$

коэффициенты a_{ij} и правые части b_i которой определяются равенствами (8). Заметим, что

$$a_{ii} = \binom{m+1-i}{m-i} = m+1-i \neq 0 \qquad (i=1, \ldots, m).$$

В пункте 2.1 было доказано, что при этом условии решение $x_1, \ldots, x_m$ системы (9) определяется равенствами

$$x_{n+1} = \frac{1}{a_{n+1,n+1}}\left[b_{n+1} - \sum_{j=1}^{n} a_{n+1,j}x_j\right] \qquad (n=0, \ldots, m-1).$$

(Если $n=0$, то правая часть равна b_1/a_{11}.) Используя (8), получаем равенства:

$$\left.\begin{aligned} x_j = a_{m+1-j},\ x_{n+1} = a_{m-n}, \\ a_{n+1,j} = \binom{m+1-j}{m-1-n},\quad a_{n+1,n+1} = \binom{m-n}{m-n-1} = m-n, \\ b_{n+1} = -\frac{1}{m+1}\binom{m+1}{m-1-n}. \end{aligned}\right\}$$

Поэтому в старых обозначениях решение системы (9) имеет вид:

$$a_{m-n} = -\frac{1}{m-n}\left[\frac{1}{m+1}\binom{m+1}{m-1-n} + \sum_{j=1}^{n}\binom{m+1-j}{m-1-n}a_{m+1-j}\right]$$
$$(n=0, \ldots, m-1). \tag{10}$$

Так как система (7) эквивалентна системе (9), коэффициенты и правые части которой определяются равенствами (8), то равенства (10) определяют решение $a_m, \ldots, a_1$ системы (7). Тем самым задача о сумме степеней решена. Это и выражает следующая основная

Теорема. *Для полинома* $p(x) = \frac{1}{m+1}x^{m+1} + a_m x^m + \ldots + a_1 x$, *коэффициенты* $a_m, \ldots, a_1$ *которого определяются равенствами* (*10*), *верны равенства*

$$p(n) = \sum_{k=0}^{n-1} k^m \qquad (n=1, 2, 3, \ldots).$$

Доказательство. 1. В пункте 3.2 было показано, что равенства

$$p(n) = \sum_{k=0}^{n-1} k^m \quad (n = 1, 2, 3, \ldots) \tag{*}$$

будут верны, если верны равенства

$$\Delta p(n) = n^m \quad (n = 0, 1, 2, \ldots). \tag{**}$$

2. В пункте 1.3 было показано, что равенства (**), в свою очередь, будут верны, если положить $a_0 = 0$, $a_{m+1} = 1/(m+1)$ и определить остальные коэффициенты $a_m, \ldots, a_1$ так, чтобы были верны равенства (7).

3. Наконец, в данном пункте показано, что равенства (7) верны, если верны равенства (10).

Следовательно, для полинома $p(x)$, коэффициенты которого определяются равенствами (10), верны равенства (*), что и требовалось доказать.

3.3. Примеры

Используя равенства (10), нетрудно получить простые выражения для нескольких старших коэффициентов.

В частности, для каждого номера $m > 0$ получаем:

$$a_m = -\frac{1}{m}\cdot\frac{1}{m+1}\binom{m+1}{m-1} = -\frac{1}{m(m+1)}\cdot\frac{(m+1)!}{(m-1)!2} =$$

$$= -\frac{(m+1)m}{m(m+1)2} = -\frac{1}{2}.$$

1. Если $m > 1$, то

$$a_{m-1} = -\frac{1}{m-1}\left[\frac{1}{m+1}\binom{m+1}{m-2} - \frac{1}{2}\binom{m}{m-2}\right] =$$

$$= -\frac{1}{m-1}\left[\frac{1}{m+1}\cdot\frac{(m+1)!}{(m-2)!3!} - \frac{1}{2}\cdot\frac{m!}{(m-2)!2!}\right] =$$

$$= -\frac{1}{m-1}\left[\frac{m(m-1)}{6} - \frac{m(m-1)}{4}\right] = \frac{m}{12}.$$

2. Если $m > 2$, то

$$a_{m-2} = -\frac{1}{m-2}\left[\frac{1}{m+1}\binom{m+1}{m-3} - \frac{1}{2}\binom{m}{m-3} + \frac{m}{12}\binom{m-1}{m-3}\right] =$$

$$= -\frac{1}{m-2}\left[\frac{1}{m+1}\cdot\frac{(m+1)!}{(m-3)!4!} - \frac{1}{2}\cdot\frac{m!}{(m-3)!4!} + \frac{m}{12}\cdot\frac{(m-1)!}{(m-3)!2!}\right] =$$

$$= -\frac{1}{m-2}\cdot\frac{m!}{(m-3)!2!}\left[\frac{1}{m+1}\cdot\frac{m+1}{12} - \frac{1}{6} + \frac{1}{12}\right] = 0.$$

3. Если $m>3$, то

$$a_{m-3}=-\frac{1}{m-3}\left[\frac{1}{m+1}\binom{m+1}{m-4}-\frac{1}{2}\binom{m}{m-4}+\frac{m}{12}\binom{m-1}{m-4}\right]=$$

$$=-\frac{1}{m-3}\left[\frac{1}{m+1}\cdot\frac{(m+1)!}{(m-4)!5!}-\frac{1}{2}\frac{m!}{(m-4)!4!}+\frac{m}{12}\cdot\frac{(m-1)!}{(m-4)!3!}\right]=$$

$$=-\frac{m!}{(m-3)!3!}\cdot\frac{1}{120}=-\frac{1}{120}\binom{m}{3}.$$

Полученные результаты можно записать в виде следующей таблицы:

a_{m+1}	a_m	a_{m-1}	a_{m-2}	a_{m-3}
$\frac{1}{m+1}$	$-\frac{1}{2}$	$\frac{m}{12}$	0	$-\frac{1}{120}\binom{m}{3}$

Пользуясь этой таблицей, легко выписать представления для суммы

$$\sum_{k=0}^{n=1} k^m$$

при $m=1, 2, 3, 4$. (В случаях $m=1, 2, 3$ результаты, естественно, совпадают с полученными в пункте 1.2.)

Имеем ($n=1, 2, 3, \ldots$):

$$\sum_{k=0}^{n-1} k=\frac{1}{2}n^2-\frac{1}{2}n,$$

$$\sum_{k=0}^{n-1} k^2=\frac{1}{3}n^3-\frac{1}{2}n^2+\frac{1}{6}n,$$

$$\sum_{k=0}^{n-1} k^3=\frac{1}{4}n^4-\frac{1}{2}n^3+\frac{1}{4}n^2,$$

$$\sum_{k=0}^{n-1} k^4=\frac{1}{5}n^5-\frac{1}{2}n^4+\frac{1}{3}n^3-\frac{1}{30}n.$$

Упражнение. Для произвольного $m>4$ вычислить коэффициент a_{m-4} и получить представление для суммы

$$\sum_{k=0}^{n-1} k^5.$$

Замечание. Задача о сумме степеней является частным случаем общей задачи о суммировании функций. Постановка и решение этой общей задачи подробно изложены в книге А. О. Гельфонда «Исчисление конечных разностей» (М., Гостехиздат, 1959). Там

рассматривается и задача о сумме степеней. Доказывается, что для каждых номеров m и $l \leqslant m$ верно равенство

$$a_{m+1-l} = \frac{1}{m+1} \binom{m+1}{l} B_l,$$

где B_l — числа Бернулли, и приводится таблица для чисел Бернулли с номерами $l \leqslant 35$. В частности,

l	0	1	2	4	6	8	10	12	14	16	18	20
B_l	1	$-\frac{1}{2}$	$\frac{1}{6}$	$-\frac{1}{30}$	$\frac{1}{42}$	$-\frac{1}{30}$	$\frac{5}{66}$	$-\frac{691}{2730}$	$\frac{7}{6}$	$-\frac{3617}{510}$	$\frac{43867}{798}$	$-\frac{174611}{330}$

(Все числа Бернулли с нечетными номерами, кроме B_1, равны нулю.)

Используя указанное равенство и таблицу, можно легко получить ответ для задачи о сумме пятых степеней, предложенной в качестве упражнения. В свою очередь, используя равенства (10), можно проверить данную таблицу.

Раздел II
ВЕРОЯТНОСТЬ

Какова вероятность того, что произойдет данное событие?

К этому вопросу сводится большое количество самых разных задач:

Какова вероятность того, что монета упадет гербом вверх?

Какова вероятность того, что стрелок попадет в цель?

Какова вероятность того, что наугад выбранное изделие окажется бракованным?

Какова вероятность того, что родится мальчик?

Теория вероятностей изучает способы вычисления вероятностей случайных событий. Она формулирует правила подсчета вероятностей некоторых простых событий. Удачное использование этих правил позволяет вычислять вероятности более сложных событий.

Теория вероятностей изучает также способы вычисления средних значений случайных переменных.

Вероятностные расчеты используются сейчас почти во всех областях науки, техники и производства.

Часть I

КОНЕЧНЫЕ ВЕРОЯТНОСТНЫЕ МОДЕЛИ

Мир, в котором мы живем, очень сложен. Присущие каждому явлению закономерности прокладывают себе путь сквозь хаос случайностей. Для многих явлений влияние случая настолько существенно, что их изучение и практическое использование невозможны без исследования и количественной оценки этого влияния.

Теория вероятностей рассматривает математические модели явлений, позволяющие учитывать влияние случая.

Глава 1

МОДЕЛЬ ЛАПЛАСА

Классическое использование термина *вероятность* связано с представлением об опытах с равновозможными результатами. Вероятностью события в такой модели называется отношение числа составляющих это событие исходов к числу всех возможных исходов. Это определение было предложено Лапласом в начале прошлого века.

§ 1. ПРИМЕРЫ

Рассмотрим несколько примеров, разъясняющих содержание модели Лапласа и подготавливающих соответствующие формальные определения.

1.1. Пример 1. Игральная кость

Игральная кость представляет собой кубик, грани которого отмечены номерами 1, 2, 3, 4, 5, 6, обозначающими числа очков.

Рассмотрим опыт: *игральная кость подбрасывается один раз.* Построим вероятностную модель этого опыта.

1.1.1. Множество исходов

Первая задача, возникающая при создании вероятностной модели рассматриваемого опыта, состоит в математическом описании его возможных *результатов*. Будем считать, что возможен один и только один из следующих результатов: кость падает гранью номер 1 вверх (*1 очко*), кость падает гранью номер 2 вверх (*2 очка*), ..., кость падает гранью номер 6 вверх (*6 очков*). Все другие результаты (например, падение кости на ребро) считаются невозможными. Каждый возможный результат удобно обозначать коротко соответствующим числом очков. Условимся эти числа называть *исходами*. Составляемое этими числами множество исходов $U = \{1, 2, 3, 4, 5, 6\}$ является удобным математическим описанием возможных результатов опыта с подбрасыванием игральной кости.

2.1.1. Алгебра событий

Вторая задача, возникающая при создании вероятностной модели рассматриваемого опыта, состоит в математическом описании *явлений*, определяемых его возможными результатами. Так как возможные результаты опыта описываются исходами, то определяемые этими результатами явления естественно описывать множествами, составленными из исходов. Например: появление четного числа очков описывается множеством $A = \{2, 4, 6\}$, появление нечетного числа очков — множеством $B = \{1, 3, 5\}$, появление числа очков, строго меньшего пяти,— множеством $C = \{1, 2, 3, 4\}$.

Такое описание учитывает логическую связь явлений: реализация *хотя бы одного* из двух явлений описывается объединением соответствующих множеств исходов; реализация *каждого* из двух явлений — *пересечением* соответствующих множеств исходов; *нереализация* явления — *дополнением* соответствующего множества исходов до множества всех исходов. Например: появление четного или нечетного числа очков описывается объединением $A \cup B = \{1, 2, 3, 4, 5, 6\}$ множеств $A = \{2, 4, 6\}$ и $B = \{1, 3, 5\}$, появление четного числа очков, строго меньшего пяти,— пересечением $A \cap C = \{2, 4\}$ множеств $A = \{2, 4, 6\}$ и $C = \{1, 2, 3, 4\}$, непоявление четного числа очков — множеством $A' = \{1, 3, 5\}$.

Каждое множество, составленное из исходов, условимся называть *событием*. Класс всех таких множеств с определенными для него объединением, пересечением и дополнением образует алгебру событий $\mathscr{A}$. Алгебра событий $\mathscr{A}$ служит удобным математическим описанием явлений, определяемых возможными результатами опыта с подбрасыванием игральной кости.

3.1.1. Элементарная вероятность

Третья задача, возникающая при создании вероятностной модели рассматриваемого опыта, состоит в математическом описании

зависимости от случая его возможных результатов, в математическом определении *меры реализуемости* этих *результатов*.

Симметричность игральной кости позволяет считать все результаты опыта с подбрасыванием кости *равновозможными*: одинаково возможно получить 1 очко, 2 очка, ..., 6 очков. Поэтому разумно приписать этим результатам *равные* меры реализации. Для того, чтобы иметь возможность сравнивать опыты с различными числами результатов, разумно выбрать для меры реализуемости значения, отнесенные к числу возможных результатов. Из сказанного следует, что каждому результату опыта с подбрасыванием игральной кости можно приписать меру реализуемости, равную 1/6. Математически эта мера определяется следующим образом. Для каждого исхода u определим число $p(u)$ равенством

$$p(u)=1/n(U).$$

Более подробно это определение выражается равенствами

$$p(1)=p(2)=p(3)=p(4)=p(5)=p(6)=1/6$$

или таблицей

u	1	2	3	4	5	6
$p(u)$	1/6	1/6	1/6	1/6	1/6	1/6

Тем самым на множестве исходов U определена функция p, значение которой для каждого исхода u равно числу $p(u)$. Назовем эту функцию p *элементарной вероятностью*. Ее значение $p(u)$ условимся называть элементарной вероятностью исхода u. Элементарная вероятность p является удобным математическим описанием меры реализуемости возможных результатов опыта с подбрасыванием игральной кости.

4.1.1. Вероятность

Четвертая задача, возникающая при создании вероятностной модели рассматриваемого опыта, состоит в математическом описании зависимости от случая явлений, определяемых возможными результатами опыта, в математическом определении *меры реализуемости* этих *явлений*.

Каждому явлению, определяемому возможными результатами рассматриваемого опыта, естественно в качестве меры реализуемости приписать сумму мер тех результатов опыта, при которых это явление реализуется. Например: появлению четного числа очков естественно приписать меру реализуемости $1/6+1/6+1/6=$ $=3/6=1/2$; появлению нечетного числа очков — ту же меру; появлению числа очков, строго меньшему пяти,— меру реализуемости $1/6+1/6+1/6+1/6=4/6=2/3$. Математически такая мера опреде-

ляется следующим образом. Для каждого события A определим число $P(A)$ равенством

$$P(A) = \sum_{u \in A} p(u),$$

т. е. число $P(A)$ равно сумме чисел $p(u)$ для исходов u, составляющих событие A. Так как все слагаемые $p(u)$ равны числу $1/n(U)$, то сумма $\sum_{u \in A} p(u)$ равна произведению числа $1/n(U)$ на число $n(A)$ слагаемых. Значит,

$$P(A) = n(A)/n(U).$$

Например:

$$P(\{2, 4, 6\}) = p(2) + p(4) + p(6) = 3/6 = 1/2,$$

$$P(\{1, 3, 5\}) = p(1) + p(3) + p(5) = 3/6 = 1/2,$$

$$P(\{1, 2, 3, 4\}) = p(1) + p(2) + p(3) + p(4) = 4/6 = 2/3.$$

Таким образом, на алгебре событий $\mathcal{A}$ определена функция P, значение которой для каждого события A равно числу $P(A)$. Назовем эту функцию P *вероятностью*. Ее значение $P(A)$ условимся называть вероятностью события A. Эта вероятность равна отношению числа $n(A)$ исходов, составляющих событие A, к числу $n(U)$ всех исходов. Вероятность P является удобным математическим описанием меры реализуемости явлений, определяемых возможными результатами опыта с подбрасыванием кости.

5.1.1. Вероятностная модель

Множество исходов U, алгебра событий $\mathcal{A}$, элементарная вероятность p и вероятность P образуют удобную модель опыта с подбрасыванием игральной кости. Термины *исход* и *событие* используются в этой модели в соответствии с их содержанием. Вероятность играет роль меры реализуемости, что также соответствует обычному смыслу этого слова. Поэтому слово *вероятность* используется и при содержательных формулировках задач.

С помощью построенной вероятностной модели решается, например, такая

Задача. *Какова вероятность того, что при подбрасывании игральной кости появляется четное число очков?*

Решение. Задача сводится к вычислению вероятности $P(A)$ события $A = \{2, 4, 6\}$. Имеем:

$$P(A) = n(A)/n(U) = 3/6 = 1/2.$$

2.1. Пример 2. Урна с шаром

В урне находится 10 шаров: 4 красных и 6 белых. Рассмотрим опыт: *из урны выбираются наугад 3 шара.* Построим вероятностную модель этого опыта.

1.2.1. Множество исходов

Обозначим красные шары номерами 0, 1, 2, 3, а белые — номерами 4, 5, 6, 7, 8. 9. Каждый возможный результат опыта можно тогда обозначить множеством номеров трех выбираемых шаров. Например: выбор шаров с номерами 0, 2, 5 обозначается множеством $\{0, 2, 5\}$, выбор шаров с номерами 7, 8, 9 — множеством $\{7, 8, 9\}$. Условимся множества, составленные из трех цифр, называть *исходами*. Составляемое ими множество исходов U является удобным математическим описанием возможных результатов опыта с выбором шаров.

2.2.1. Алгебра событий

Так как возможные результаты опыта описываются исходами, то определяемые этими результатами явления естественно описывать множествами, составленными из исходов. Например, появление только красных шаров описывается множеством $A=\{\{0, 1, 2\}, \{0, 1, 3\}, \{0, 2, 3\}, \{1, 2, 3\}\}$. Как и в примере с игральной костью, такое описание учитывает логическую связь явлений.

Каждое множество, составленное из исходов, условимся называть *событием*. Класс всех таких множеств с определенными для него объединением, пересечением и дополнением образует алгебру событий $\mathscr{A}$. Алгебра событий $\mathscr{A}$ является удобным математическим описанием явлений, определяемых возможными результатами опыта с выбором шаров.

3.2.1. Элементарная вероятность

Выбор наугад шаров из урны позволяет считать все результаты опыта равновозможными: одинаково возможно выбрать любые три шара. Поэтому разумно приписать этим результатам меру реализуемости, равную $1\Big/\binom{10}{3}$. Математически эта мера определяется как функция p на множестве исходов U, значение которой для каждого исхода u равно числу

$$p(u) = 1/n(U) = 1\Big/\binom{10}{3} = 1/120.$$

Назовем эту функцию *p элементарной вероятностью*. Ее значение $p(u)$ условимся называть элементарной вероятностью исхода u. Элементарная вероятность p является удобным математическим определением меры реализуемости возможных результатов опыта с выбором шаров.

4.2.1. Вероятность

Каждому явлению, определяемому возможными результатами рассматриваемого опыта, естественно в качестве меры реализуемо-

сти приписать сумму мер тех результатов опыта, при которых это явление реализуется. Математически эта мера определяется как функция P на алгебре событий $\mathscr{A}$, значение которой для каждого события A равно числу

$$P(A) = \sum_{u \in A} p(u) = n(A)/n(U).$$

Назовем функцию P *вероятностью*. Ее значение $P(A)$ условимся называть вероятностью события A. Эта вероятность равна отношению числа $n(A)$ исходов, составляющих событие A, к числу $n(U)$ всех исходов. Например, вероятность $P(A)$ события $A = \{\{0, 1, 2\}, \{0, 1, 3\}, \{0, 2, 3\}, \{1, 2, 3\}\}$, описывающего появление только красных шаров, выражается равенствами

$$P(A) = 1/120 + 1/120 + 1/120 + 1/120 = 4/120.$$

Вероятность P является удобным математическим описанием меры реализуемости явлений, определяемых возможными результатами опыта с выбором шаров.

5.2.1. Вероятностная модель

Множество исходов U, алгебра событий $\mathscr{A}$, элементарная вероятность p и вероятность P образуют удобную модель рассматриваемого опыта с выбором шаров. Термины *исход*, *событие* и *вероятность* используются в ней в соответствии с их содержанием. Поэтому они употребляются и при содержательных описаниях задач.

С помощью построенной вероятностной модели решается, например, такая

Задача. *Какова вероятность того, что среди выбираемых 3 шаров ровно 2 — красные?*

Решение. Задача сводится к вычислению вероятности $P(A)$ события A, составленного из трех исходов, в которых два номера меньше 3, а один — строго больше 3. Событию A принадлежат, например, исходы $\{0, 1, 4\}$, $\{0, 1, 5\}$, $\{0, 2, 4\}$, $\{2, 3, 9\}$. Вычисление вероятности $P(A)$ сводится к подсчету числа $n(A)$ исходов, составляющих событие A. Используя теорему о числе выборок и правило умножения, нетрудно убедиться в том, что

$$n(A) = \binom{4}{2}\binom{6}{1} = 6 \cdot 6 = 36,$$

откуда

$$P(A) = n(A)/n(U) = 36/120 = 3/10$$

3.1. Пример 3. Игра с монетой

Рассмотрим такой опыт: *симметричная монета подбрасывается два раза.* С этим опытом можно связать игру:

1) если монета *оба раза* падает *гербом* вверх, то игрок *выигрывает*;

2) если монета *оба раза* падает *цифрой* вверх, то игрок *проигрывает*;

3) если монета *один раз* падает *гербом* вверх и *один — цифрой*, то игра оканчивается *ничьей* (игрок не выигрывает и не проигрывает).

Построим две различные вероятностные модели такой игры.

1.3.1. Первая модель

Будем рассматривать то, что интересует игрока: выигрыш, ничья и проигрыш.

1.1.3.1. Множество исходов. Возможными результатами опыта условимся считать следующие:

0) *проигрыш* (монета ни разу не падает гербом вверх);

1) *ничья* (монета ровно один раз падает гербом вверх);

2) *выигрыш* (монета два раза падает гербом вверх).

Номера $u_1=0, 1, 2$ этих результатов назовем *исходами*. Множество исходов $U_1=\{0, 1, 2\}$ описывает возможные результаты опыта.

2.1.3.1. Алгебра событий. Явления, определяемые возможными результатами опыта, будем описывать частями множества исходов U_1. Назовем эти части *событиями*. Например: выигрыш описывает событие $B_1=\{2\}$, проигрыш — событие $C_1=\{0\}$, ничья — событие $D_1=\{1\}$, непроигрыш (монета хотя бы один раз падает гербом вверх) — событие $A_1=\{1, 2\}$.

Класс

$$\mathscr{A}_1=\{O, \{0\}, \{1\}, \{2\}, \{0, 1\}, \{0, 2\}, \{1, 2\}, U\}$$

всех частей множества U_1 с определенными для него объединением, пересечением и дополнением образует алгебру событий $(\mathscr{A}_1, \cup, \cap, ')$.

3.1.3.1. Элементарная вероятность. Меру реализуемости возможных результатов опыта определим, считая все результаты равновозможными. Припишем каждому исходу $u_1=0, 1, 2$ элементарную вероятность $p_1(u_1)=1/3$. Функцию p_1 *на множестве исходов* U_1 со значениями $p_1(0)=p_1(1)=p_1(2)=1/3$ назовем элементарной вероятностью.

4.1.3.1. Вероятность. Меру реализуемости явлений определяемых возможными результатами опыта, введем следующим образом. Каждому событию A_1 припишем вероятность $P_1(A_1)$, равную сумме элементарных вероятностей $p_1(u_1)$ исходов u_1, составляющих событие A_1:

$$P(O)=0;$$

$$P_1(\{0\})=p(0)=1/3,\ P(\{1\})=p(1)=1/3,\ P(\{2\})=p(2)=1/3;$$

$$P_1(\{0, 1\})=p(0)+p(1)=2/3,\quad P(\{0, 2\})=p(0)+p(2)=2/3,$$

$$P(\{1, 2\}) = p(1) + p(2) = 2/3;$$

$$P_1(U_1) = p(0) + p(1) + p(2) = 1.$$

Функцию P_1 на классе событий $\mathscr{A}_1$ со значением $P_1(A_1)$ для каждого события A_1 назовем *вероятностью.*

5.1.3.1. Вероятностная модель 1. Множество исходов U_1, алгебра событий $\mathscr{A}_1$, элементарная вероятность p_1 и вероятность P_1 образуют первую модель игры с монетой.

В этой модели вероятности событий B_1, C_1, D_1 и A_1, описывающих выигрыш, проигрыш, ничью и непроигрыш (появление герба два раза, нуль раз, ровно один раз и хотя бы один раз), соответственно равны:

$$P_1(B_1) = P_1(C_1) = P_1(D_1) = 1/3, \; P_1(A_1) = 2/3.$$

2.3.1. Вторая модель

Рассмотрим опыт с подбрасываниями монеты подробнее.

1.2.3.1. Множество исходов. Возможными результатами опыта условимся считать:

00) *первый* раз появляется *цифра, второй* раз появляется *цифра*;

01) *первый* раз появляется *цифра, второй* раз появляется *герб*;

10) *первый* раз появляется *герб, второй* раз появляется *цифра*;

11) *первый* раз появляется *герб, второй* раз появляется *герб.*

Для описания этих результатов используем пары 00, 01, 10, 11. Назовем эти пары *исходами.* Множество исходов $U_2 = \{00, 01, 10, 11\}$ описывает возможные результаты опыта.

2.2.3.1. Алгебра событий. Явления, определяемые возможными результатами опыта, будем описывать частями множества исходов U_2. Назовем эти части *событиями.* Например:

появление герба два раза (*выигрыш*) описывает событие $B_2 = \{11\}$,

появление герба нуль раз (*проигрыш*) — событие $C_2 = \{00\}$,

появление герба ровно один раз (*ничья*) — $D_2 = \{01, 10\}$,

появление герба хотя бы один раз (*непроигрыш*) — событие $A_2 = B_2' = \{01, 10, 11\}$.

Класс $\mathscr{A}_2$ всех частей множества U_2 с определенными для него объединением, пересечением и дополнением образует алгебру событий $(\mathscr{A}_2, \bigcup, \bigcap, ')$.

3.2.3.1. Элементарная вероятность. Меру реализуемости возможных результатов опыта определим, считая все результаты равновозможными. Припишем каждому исходу $u_2 = 00, 01, 10, 11$ элементарную вероятность $p_2(u_2) = 1/4$. Функцию p_2 на множестве исходов U_2 со значением $p_2(00) = p(01) = p(10) = p(11) = 1/4$ назовем *элементарной вероятностью.*

4.2.3.1. Вероятность. Меру реализуемости явлений, определяемых возможными результатами опыта, введем, приписав каждому

событию A_2 вероятность $P_2(A_2)$, равную сумме элементарных вероятностей $p_2(u_2)$ исходов u_2, составляющих событие A_2. Функцию P_2 на классе событий $\mathscr{A}_2$ со значением $P_2(A_2)$ для каждого события A_2 назовем *вероятностью*.

5.2.3.1. Вероятностная модель 2. Множество исходов U_2, алгебра событий $\mathscr{A}_2$, элементарная вероятность p_2 и вероятность P_2 образуют вторую модель игры с монетой.

В этой модели вероятности событий B_2, C_2, D_2 и A_2, описывающих появление герба два раза, нуль раз, ровно один раз и хотя бы один раз (выигрыш, проигрыш, ничья и непроигрыш), соответственно равны:

$$P_2(B_2) = P_2(C_2) = 1/4, \quad P_2(D_2) = 2/4, \quad P_2(A_2) = 3/4.$$

3.3.1. Обсуждение

Таким образом, различные модели дали различные вероятности событий, описывающих одни и те же явления. Возникает вопрос, какая из предлагаемых моделей больше соответствует действительности? Анализируя их, можно заметить, что в первой модели предположение о равновозможности результатов ничем не обосновано, а во второй оно обосновывается симметричностью монеты. Это замечание позволяет думать, что вторая модель больше соответствует действительности. Однако окончательно решить этот вопрос можно только практической проверкой. Если вероятность события действительно является мерой реализуемости соответствующего явления, то она должна быть близка частоте, с которой это явление происходит при реальных подбрасываниях монеты, повторенных достаточно большое число раз.

Практическую проверку соответствия рассматриваемых вероятностных моделей действительности провести нетрудно. В частности, когда монета дважды подбрасывалась 50 раз, то результат «цифра, цифра» наблюдался 12 раз, результат «цифра, герб» — 14 раз, «герб, цифра» — 12 раз, «герб, герб» — 12 раз. Следовательно, результат «герб ноль раз» наблюдался 12 раз, «герб ровно один раз» — 26 раз, «герб два раза» — 12 раз. Соответствующие частоты выражаются таблицами:

0	1	2
$\frac{12}{50}=\frac{1}{3}-\frac{1}{75}$	$\frac{26}{50}=\frac{1}{3}+\frac{14}{75}$	$\frac{12}{50}=\frac{1}{3}-\frac{1}{75}$

;

00	01	10	11
$\frac{12}{50}=\frac{1}{4}-\frac{1}{100}$	$\frac{14}{50}=\frac{1}{4}+\frac{3}{100}$	$\frac{12}{50}=\frac{1}{4}-\frac{1}{100}$	$\frac{12}{50}=\frac{1}{4}-\frac{1}{100}$

.

Эти таблицы показывают, что во второй модели наблюденные частоты ближе соответствующим элементарным вероятностям. Проделанный эксперимент подтверждает большее соответствие действительности второй модели.

Упражнение. Подбросить дважды монету 100 раз и сравнить наблюденные частоты с элементарными вероятностями первой и второй моделей.

§ 2. МОДЕЛЬ ЛАПЛАСА

Для описания опытов с *равновозможными* результатами удобна классическая модель Лапласа.

1.2. Множество исходов U

Основой модели является непустое конечное множество U. Элементы u множества U называются *исходами*. Само множество U называется *множеством исходов*. Множество исходов U описывает *возможные результаты* опыта.

2.2. Алгебра событий $\mathscr{A}$

Вторым элементом модели является алгебра $(\mathscr{A}, \bigcup, \bigcap, ')$, образованная классом $\mathscr{A}$ всех частей множества исходов U и определенными для него объединением, пересечением и дополнением. Части A множества исходов U называются *событиями*. Алгебра $(\mathscr{A}, \bigcup, \bigcap, ')$ называется алгеброй событий и сокращенно обозначается буквой $\mathscr{A}$ (так же, как и класс событий). Алгебра событий $\mathscr{A}$ описывает *явления, определяемые возможными результатами опыта.*

Пустое множество O называется *невозможным* событием, а множество исходов U — *достоверным*; каждое событие $A \neq O$ называется *возможным*.

3.2. Элементарная вероятность p

Третьим элементом модели является элементарная вероятность p.

Рассмотрим функцию p на множестве исходов U, значение $p(u)$ которой для каждого исхода u определяется равенством

$$p(u) = 1/n(U). \tag{1}$$

Функция p называется *элементарной вероятностью*. Ее значение $p(u)$ — *элементарной вероятностью исхода u*. Элементарная вероятность описывает зависимость возможных результатов опыта от случая. Она является *мерой реализуемости* этих *результатов* Равенство (1) описывает их равновозможность.

4.2. Вероятность P

Четвертым и самым важным элементом модели является вероятность P.

Для каждого события A рассмотрим сумму $P(A)$ элементарных вероятностей $p(u)$ исходов u, составляющих событие A:

$$P(A) = \sum_{u \in A} p(u) = n(A)/n(U). \tag{2}$$

Эти равенства определяют на классе событий $\mathscr{A}$ функцию P со значением $P(A)$ для каждого события A. Функция P называется *вероятностью*, а ее значение $P(A)$ — ***вероятностью события** A*. Вероятность описывает зависимость от случая явлений, определяемых результатами опыта. Она служит *мерой реализуемости* этих *явлений.*

5.2. Вероятностная модель Лапласа

Множество исходов U, алгебра событий $\mathscr{A}$, элементарная вероятность p и вероятность P, определяемая равенствами (1) и (2), образуют классическую **модель Лапласа.**

Замечание. Конкретный выбор множества исходов U определяется удобством описания рассматриваемых явлений. Существенным для модели является только число $n(U)$ элементов множества U. Можно нумеровать каким-либо образом результаты опыта и использовать для его описания *стандартную* модель Лапласа n равновероятных исходов $1, \ldots, n$, как это сделано в примерах 1 и 2.

Подчеркнем еще раз, что модель Лапласа описывает опыты с *равновозможными* результатами. Пример 3 из § 1 показывает, что если результаты не равновозможны, то эта модель не применима.

§ 3. СВОЙСТВА ВЕРОЯТНОСТИ

Рассмотрим основные свойства элементарной вероятности и вероятности для модели Лапласа.

1.3. Элементарная вероятность p

Элементарная вероятность p является положительной нормированной функцией на множестве исходов U.

1. Элементарная вероятность p *положительна*:

$$p(u) = 1/n(U) \geqslant 0$$

для каждого исхода u.

2. Элементарная вероятность p *нормирована.* Это значит, что сумма элементарных вероятностей $p(u)$ для всех исходов u равна единице:

$$\sum_{u \in U} p(u) = n(U) \cdot 1/n(U) = 1.$$

2.3. Вероятность P

Вероятность P является положительной, нормированной и аддитивной функцией на классе событий $\mathscr{A}$.

1. Вероятность P *положительна*:

$$P(A) = \sum_{u \in A} = n(A)/n(U) \geqslant 0$$

для каждого события A.

2. Вероятность P *нормирована*:

$$P(U) = \sum_{u \in U} p(u) = n(U)/n(U) = 1.$$

3. Вероятность P *аддитивна*:

$$P(A+B) = P(A) + P(B)$$

для каждых непересекающихся событий A и B. Это правило сложения для вероятности P в модели Лапласа может быть получено как следствие правила сложения для числа элементов:

$$P(A+B) = n(A+B)/n(U) =$$
$$= (n(A) + n(B))/n(U) = n(A)/n(U) + n(B)/n(U) =$$
$$= P(A) + P(B).$$

3.3. Примеры

Рассмотрим простые примеры использования правила сложения для вычисления вероятностей событий.

Пример 1. *Игральная кость подбрасывается один раз. Какова вероятность того, что появится четное или нечетное число очков?*

Решим эту задачу, используя модель примера 1 из § 1. Она сводится к вычислению вероятности $P(A+B)$ суммы $A+B$ непересекающихся событий $A = \{2, 4, 6\}$ и $B = \{1, 3, 5\}$. Используя правило сложения, находим:

$$P(A+B) = P(A) + P(B) = 3/6 + 3/6 = 1.$$

Тот же результат получается, если использовать равенство $A+B=U$:

$$P(A+B) = P(U) = 6/6 = 1.$$

Пример 2. *В урне находятся 4 красных шара и 6 белых. Из этих 10 шаров наугад выбираются 3 шара. Какова вероятность того, что все выбранные шары одного цвета?*

Решим эту задачу, используя модель примера 2 из § 1. Она сводится к вычислению вероятности $P(A+B)$ суммы $A+B$ непересекающихся событий A и B, описывающих соответственно появление 3 красных шаров и 3 белых. По теореме о числе выборок

$$n(A) = \binom{4}{3} = 4, \; n(B) = \binom{6}{3} = 20, \; n(U) = \binom{10}{3} = 120.$$

Используя правило сложения, получаем:

$$P(A+B)=P(A)+P(B)=4/120+20/120=1/5.$$

Пример 3. *Симметричная монета подбрасывается два раза. Какова вероятность того, что она хотя бы один раз падает гербом вверх?*

Решим эту задачу, используя вторую модель примера 3 из § 1. Она сводится к вычислению вероятности $P(B+C)$ суммы $B+C$ непересекающихся событий $B=\{11\}$ и $C=\{01, 10\}$. Используя правило сложения, получаем:

$$P(B+C)=P(B)+P(C)=1/4+2/4=3/4.$$

§ 4. УСЛОВНАЯ ВЕРОЯТНОСТЬ

При оценке реализуемости явления часто приходится учитывать различные дополнительные условия. Для этого в рассматриваемой модели определяется *условная вероятность*.

1.4. Примеры

Рассмотрим несколько примеров, разъясняющих содержание понятия условной вероятности.

Пример 1. *Игральная кость подбрасывается один раз. Известно, что появляется нечетное число очков. Какова вероятность при этом условии того, что появляется число очков, строго меньшее четырех?*

Для наглядности представим себе, что грани с нечетными числами очков имеют красный цвет, а грани с четными числами очков — белый. Алик накрывает подброшенную кость ладонью и предлагает Боре принять участие в следующей игре: если появившееся число очков строго меньше четырех, то Боря выигрывает пять копеек, если больше — проигрывает. Боря успел заметить, что кость упала красной гранью вверх. Каковы его шансы на выигрыш в этом случае?

Решим поставленную задачу, используя модель примера 1 из § 1. Рассмотрим те же множество исходов U и алгебру событий $\mathcal{A}$, что и в этой модели. Они по-прежнему хорошо описывают опыт с подбрасыванием игральной кости. В то же время ясно, что для того, чтобы учесть условие о появлении нечетного числа очков, вероятность нужно определить не так, как это было сделано в примере 1 из § 1.

Нечетное число очков получается в результате появления 1, 3 или 5 очков. Симметричность кости по-прежнему позволяет считать эти результаты равновозможными. Условие о появлении нечетного числа очков заставляет считать появление 2, 4 или 6 очков нереализуемыми результатами. Эти соображения приводят к следующему определению новой, условной вероятности. Рассмотрим событие

$B=\{1, 3, 5\}$, описывающее появление нечетного числа очков. Определим *условную элементарную вероятность* p_B равенствами

$$p_B(1)=p_B(3)=p_B(5)=1/n(B)=1/3,\ p(2)=p(4)=p(6)=0.$$

В игре Алика и Бори первые равенства оценивают шансы на появление 1, 3, 5 очков при условии, что кость упала красной гранью вверх. Вторые равенства выражают тот факт, что красная грань не является белой.

Для каждого события A условную вероятность $P_B(A)$ определим равенствами

$$P_B(A)=\sum_{u\in A} p_B(u)=n(AB)/n(B).$$

Эти равенства определяют на алгебре событий $\mathscr{A}$ *условную вероятность* P_B.

Множество исходов U, алгебра событий $\mathscr{A}$, условная элементарная вероятность p_B и условная вероятность P_B образуют удобную математическую модель опыта с подбрасыванием игральной кости при условии, что появляется нечетное число очков. Используя эту модель, легко решить поставленную задачу. Она сводится к вычислению условной вероятности $P_B(A)$ события $A=\{1, 2, 3\}$. Так как $AB=\{1, 3\}$, то

$$P_B(A)=n(AB)/n(B)=2/3.$$

Если Боря знает, что кость упала красной гранью вверх, то он может оценить свои шансы на выигрыш числом 2/3.

Заметим, что

$$P(A)=n(A)/n(U)=3/6=1/2<2/3=P_B(A).$$

Если Боря ничего не знает о том, какой гранью упала кость, то он поступит разумно, оценив свои шансы на выигрыш числом 1/2, меньшим 2/3.

Пример 2. *В урне находятся 4 красных шара и 6 белых. Из этих 10 шаров наугад выбираются 3 шара. Известно, что все выбираемые шары одного цвета. Какова вероятность при этом условии того, что все выбираемые шары красные?*

Решим эту задачу, используя модель примера 2 из § 1. Рассмотрим те же множество исходов U и алгебру событий $\mathscr{A}$, что и в этой модели.

Для того, чтобы учесть условие об одноцветности выбираемых шаров, определим новую, *условную* вероятность. Шары одного цвета появляются в результате каждого выбора трех красных шаров или трех белых. Выбор шаров наугад позволяет по-прежнему считать эти результаты равновозможными. Условие об одноцветности заставляет считать каждое появление шаров разного цвета нереализуемым результатом. Эти соображения приводят к следующему определению условной вероятности. Рассмотрим событие B, описывающее появление трех шаров одного цвета. Определим *условную элементарную вероятность* p_B равенствами

$$p_B(u)=1/n(B)\quad (u\in B),\quad p_B(u)=0\quad (u\notin B).$$

Например:

$$p_B(\{0,\ 1,\ 2\}) = p_B(\{7,\ 8,\ 9\}) = 1/24,$$
$$p_B(\{0,\ 1,\ 9\}) = p_B(\{0,\ 8,\ 9\}) = 0.$$

Для каждого события A *условную вероятность* P_B определим равенствами

$$P_B(A) = \sum_{u \in A} p_B(u) = n(AB)/n(B).$$

В частности, если событие A описывает появление шаров разного цвета, то $AB = 0$ и

$$P_B(A) = n(AB)/n(B) = n(O)/n(B) = 0.$$

Если $A = B$, то $AB = B$ и

$$P_B(A) = n(AB)/n(B) = n(B)/n(B) = 1.$$

Множество исходов U, алгебра событий $\mathscr{A}$, условная элементарная вероятность p_B и условная вероятность P_B образуют удобную математическую модель опыта с выбором шаров при условии, что выбираются шары одного цвета. Используя эту модель, легко решить поставленную задачу. Она сводится к вычислению условной вероятности $P_B(A)$ события $A = \{\{0,\ 1,\ 2\},\ \{0,\ 1,\ 3\},\ \{0,\ 2,\ 3\},\ \{1,\ 2,\ 3\}\}$. Так как $AB = A$, то

$$P_B(A) = n(AB)/n(B) = n(A)/n(B) = 4/24 = 1/6.$$

Заметим, что

$$P(A) = n(A)/n(U) = 4/120 = 1/30 \neq 1/6 = P_B(A).$$

Информация об одноцветности выбранных шаров изменяет оценку шансов появиться трем красным шарам. В связи с этим можно сказать, что событие A, описывающее красный цвет, *зависит* от события B, описывающего одинаковый цвет выбираемых шаров.

Пример 3. *Симметричная монета подбрасывается два раза. Известно, что первый раз появляется герб. Какова вероятность при этом условии того, что и второй раз появится герб?*

Решим эту задачу, используя модель примера 3 из § 1. Рассмотрим то же множество исходов U и алгебру событий $\mathscr{A}$, что и в этой модели.

Для того, чтобы учесть условие о появлении герба при первом подбрасывании, определим новую, *условную* вероятность. Рассуждая по аналогии с примерами 1 и 2, рассмотрим событие $B = \{10,\ 11\}$, описывающее появление герба при первом подбрасывании. Определим *условную элементарную вероятность* p_B равенствами

$$p_B(10) = p_B(11) = 1/n(B) = 1/2,\quad p_B(00) = p_B(01) = 0.$$

Первые равенства оценивают реализуемость результатов «герб, цифра» и «герб, герб» при условии, что при первом подбрасывании появился герб. Вторые равенства выражают тот факт, что если при первом подбрасывании появляется герб, то не появляется цифра.

Для каждого события A условную вероятность $P_B(A)$ определим равенствами

$$P_B(A) = \sum_{u \in A} p_B(u) = n(AB)/n(B).$$

Эти равенства определяют на алгебре событий $\mathscr{A}$ *условную вероятность* P_B.

Множество исходов U, алгебра событий $\mathscr{A}$, условная элементарная вероятность p_B и условная вероятность P_B образуют удобную математическую модель опыта с подбрасыванием монеты два раза при условии, что первый раз появляется герб. Используя эту модель, легко решить поставленную задачу. Она сводится к вычислению вероятности $P_B(A)$ события $A = \{01, 11\}$. Так как $AB = \{11\}$, то

$$P_B(A) = n(AB)/n(B) = 1/2.$$

Заметим, что

$$P(A) = n(A)/n(U) = 2/4 = 1/2 = P_B(A).$$

Это равенство соответствует тому, что в рассматриваемом опыте результат второго подбрасывания, очевидно, не зависит от результатов первого подбрасывания. Поэтому информация о появлении герба при первом подбрасывании не изменяет оценку шансов появиться гербу при втором подбрасывании. В связи с этим можно сказать, что событие A, описывающее появление герба при втором подбрасывании, *не зависит* от события B, описывающего появление герба при первом подбрасывании.

2.4. Определения

Примеры 1—3 подводят к следующему определению условной вероятности в модели Лапласа с множеством исходов U, алгеброй событий $\mathscr{A}$, элементарной вероятностью p и вероятностью P.

Рассмотрим произвольное *возможное* событие B.

По аналогии с определением элементарной вероятности p предварительно определим *условную элементарную вероятность* p_B как функцию на множестве исходов U, значение $p(u)$ которой для каждого исхода u определяется равенствами

(3) $$p_B(u) = 1/n(B) \quad (u \in B), \quad p_B(u) = 0 \quad (u \notin B).$$

В этом определении используется предполагаемая возможность события B, обеспечивающая неравенство $n(B) \neq 0$. Условная элементарная вероятность p_B является мерой реализуемости результатов рассматриваемого опыта при условии, что реализуется явление, описываемое событием B.

По аналогии с определением вероятности P определим *условную вероятность* P_B как функцию на классе событий $\mathscr{A}$, значение $P_B(A)$ которой для каждого события A определяется равенствами

(4) $$P_B(A) = \sum_{u \in A} p_B(u) = n(AB)/n(B).$$

Условная вероятность P_B является мерой реализуемости явлений, определяемых результатами рассматриваемого опыта, при условии, что реализуется явление, описываемое событием B.

Таким образом, модель Лапласа с множеством исходов U, алгеброй событий $\mathscr{A}$, элементарной вероятностью p, вероятностью P и возможное событие B для этой модели определяют новую модель с множеством исходов U, алгеброй событий $\mathscr{A}$, условной элементарной вероятностью p_B, условной вероятностью P_B.

3.4. Правило умножения

Рассмотрим модель Лапласа с множеством исходов U, алгеброй событий $\mathscr{A}$ и вероятностью P. Для каждых события A и возможного события B в этой модели верно *правило деления*:

$$P_B(A) = P(AB)/P(B). \tag{5}$$

В самом деле, используя равенства (2) и (4), определяющие вероятность и условную вероятность, получаем:

$$P_B(A) = n(AB)/n(B) = (n(AB)/n(U))/(n(B)/n(U)) = \\ = P(AB)/P(B).$$

Из равенства (5) следует равенство

$$P(AB) = P(B)P_B(A). \tag{6}$$

Это равенство называется *правилом умножения* для вероятности.

Равенства (5) и (6) часто используются для вычисления вероятностей событий.

Пример 1. *Игральная кость подбрасывается один раз. Какова вероятность того, что появляется нечетное число очков, строго меньшее четырех?*

Решим эту задачу, используя модель примера 1 из § 1. Она сводится к вычислению вероятности $P(AB)$ пересечения AB событий $A = \{1, 2, 3\}$ и $B = \{1, 3, 5\}$. Используя правило умножения, вычисленные раньше вероятность $P(B)$ и условную вероятность $P_B(A)$, получаем:

$$P(AB) = P(B)P_B(A) = 1/2 \cdot 2/3 = 1/3.$$

В том же самом можно убедиться, непосредственно используя определение вероятности. Так как $AB = \{1, 3\}$, то

$$P(AB) = n(AB)/n(U) = 2/6 = 1/3.$$

Пример 2. *В урне находятся 4 красных шара и 6 белых. Из этих 10 шаров наугад выбираются 3 шара. Какова вероятность того, что все выбираемые шары красные?*

Решим эту задачу, используя модель примера 2 из § 1. Она сводится к вычислению вероятности $P(A)$ события A, описывающего появление только красных шаров. Заметим, что событие A рав-

но своему пересечению AB с событием B, описывающим появление шаров одного цвета. Используя правило умножения, вычисленные раньше вероятность $P(B)$ и условную вероятность $P_B(A)$, получаем:

$$P(A)=P(AB)=P(B)P_B(A)=1/5\cdot 1/6=1/30.$$

То же самое было найдено в начале параграфа при непосредственном использовании определения вероятности.

Пример 3. *Симметричная монета подбрасывается два раза. Какова вероятность того, что оба раза появляется герб?*

Решим эту задачу, используя модель примера 3 из § 1. Она сводится к вычислению вероятности $P(AB)$ пересечения AB событий $A=\{01, 11\}$ и $B=\{10, 11\}$. Используя правило умножения и проведенные раньше вычисления, получаем:

$$P(AB)=P(B)P_B(A)=P(A)P(B)=1/2\cdot 1/2=1/4.$$

То же самое получаем, непосредственно используя определение вероятности. Так как $AB=\{11\}$, то

$$P(AB)=n(AB)/n(U)=1/4.$$

4.4. Формула полной вероятности

Из правил сложения и умножения вытекает очень удобная *формула полной вероятности.* В простейшем случае эта формула имеет следующий вид.

Рассмотрим классическую модель Лапласа с множеством исходов U, алгеброй событий $\mathscr{A}$ и вероятностью P. Предположим, что события B и C не пересекаются и в сумме составляют достоверное событие U:

$$BC=O, \quad B+C=U. \tag{7}$$

В этом случае для вероятности $P(A)$ каждого события A верно равенство

$$P(A)=P(B)P_B(A)+P(C)P_C(A). \tag{8}$$

Действительно, используя правило умножения, правило сложения и равенства (7) для событий B и C, получаем:

$$P(B)P_B(A)+P(C)P_C(A)=P(AB)+P(AC)=$$
$$=P(AB+AC)=P(A(B+C))=P(AU)=P(A).$$

Равенство (8) и называется формулой полной вероятности.

Пример 1. *Игральная кость подбрасывается один раз. Какова вероятность того, что появится не шесть очков?*

Решим эту задачу, используя модель примера 1 из § 1. Она сводится к вычислению вероятности события $A=\{1, 2, 3, 4, 5\}$. Рассмотрим события $B=\{2, 4, 6\}$ и $C=\{1, 3, 5\}$. Эти события не пересекаются и в сумме составляют достоверное событие U. Для

них верны равенства (7). Используя формулу полной вероятности и проделанные раньше вычисления, находим:

$$P(A)=P(B)P_B(A)+P(C)P_C(A)=1/2\cdot 2/3+1/2\cdot 1=5/6.$$

То же самое получается при непосредственном использовании определения вероятности:

$$P(A)=n(A)/n(U)=5/6.$$

Пример 2. *В урне находятся 4 красных шара и 6 белых. Из этих 10 шаров наугад выбираются 3 шара. Какова вероятность того, что среди выбранных шаров 2 или 3 красные?*

Решим эту задачу, используя модель 2 из § 1. Она сводится к вычислению вероятности события A, описывающего появление трех шаров, среди которых два или три красные. Событие A является суммой непересекающихся событий A_2, A_3, описывающих соответственно появление среди трех выбираемых шаров ровно 2 или 3 красных.

Рассмотрим события B и C, описывающие соответственно появление неодноцветных шаров и одноцветных. Эти события не пересекаются и в сумме составляют достоверное событие U: для них верны равенства (7). Было доказано, что

$$P(C)=1/5.$$

Следовательно, по правилу разности,

$$P(B)=1-P(C)=4/5.$$

Вычислим условные вероятности, фигурирующие в формуле полной вероятности. Заметим, что $AB=A_2$ и $AC=A_3$. Отсюда и из равенства (6) следует, что

$$P_B(A)=P(AB)/P(B)=P(A_2)/P(B)=P_B(A_2),$$
$$P_C(A)=P(AC)/P(C)=P(A_3)/P(C)=P_C(A_3).$$

Было показано, что $P_C(A_3)=1/6$. Следовательно,

$$P_C(A)=1/6.$$

Было показано также, что $P(A_2)=3/10$. Следовательно,

$$P_B(A)=P(A_2)/P(B)=(3/10)/(4/5)=3/8.$$

Используя формулу полной вероятности и проделанные вычисления, получаем:

$$P(A)=P(B)P_B(A)+P(C)P_C(A)=4/5\cdot 3/8+1/5\cdot 1/6=1/3.$$

То же самое можно получить, также используя правило сложения и вычисленные раньше вероятности $P(A_2)$ и $P(A_3)$:

$$P(A)=P(A_2)+P(A_3)=3/10+1/30=1/3.$$

Пример 3. *Симметричная монета подбрасывается два раза. Какова вероятность того, что герб появляется ровно один раз?*

Решим эту задачу, используя модель примера 3 из § 1. Она сводится к вычислению вероятности события $A=\{01,\ 10\}$. Рассмотрим события $B=\{00,\ 10\}$ и $C=\{01,\ 11\}$. Эти события описывают соответственно появление цифры и герба при втором подбрасывании. Они не пересекаются и в сумме составляют достоверное событие U: для них верны равенства (7). Кроме того,

$$P(B)=1/2, \quad P(C)=1/2;$$

$$P_B(A)=n(AB)/n(B)=1/2,$$

$$P_C(A)=n(AC)/n(C)=1/2.$$

Используя формулу полной вероятности, получаем:

$$P(A)=P(B)P_B(A)+P(C)P_C(A)=1/2\cdot 1/2+1/2\cdot 1/2=1/2.$$

В том же самом можно убедиться, непосредственно используя определение вероятности:

$$P(A)=n(A)/n(U)=2/4=1/2.$$

§ 5. НЕЗАВИСИМОСТЬ И ЗАВИСИМОСТЬ

Условная вероятность позволяет определить меру зависимости между явлениями. Для этого можно использовать, например, разность между условной вероятностью и вероятностью.

1.5. Примеры

Рассмотрим несколько примеров, разъясняющих содержание вероятностной независимости и зависимости явлений.

Пример 1. В примере 1 из § 4 были вычислены вероятность $P(A)$ события $A=\{1,\ 2,\ 3\}$, описывающего появление при подбрасывании игральной кости числа очков, строго меньшего четырех, и условная вероятность $P_B(A)$ события A при условии $B=\{1, 3, 5\}$, описывающем появление нечетного числа очков:

$$P_B(A)=2/3\neq 1/2=P(A).$$

Это неравенство описывает зависимость между появлением нечетного числа очков и появлением числа очков, строго меньшего четырех. Говорят также о зависимости события A от события B. Меру этой зависимости можно выбрать разными способами. Например, использовать разность

$$P_B(A)-P(A)=2/3-1/2=1/6.$$

Рассмотрим событие $B'=\{2,\ 4,\ 6\}$, дополнительное для $B=\{1,\ 3,\ 5\}$ и описывающее появление четного числа очков. Неравенство

$$P_{B'}(A)=n(AB')/n(B')=1/3\neq 1/2=P(A)$$

описывает зависимость события A от события B'. Так как событие B' определяется событием B, то можно считать, что это неравенство также описывает зависимость события A от события B. Поэтому для измерения этой зависимости можно использовать также разность

$$P_{B'}(A) - P(A) = 1/3 - 1/2 = -1/6.$$

Естественно объединить предложенные меры зависимости и рассматривать разность

$$R(A, B) = P_B(A) - P_{B'}(A) = [P_B(A) - P(A)] - [P_{B'}(A) - P(A)] = \\ = 1/6 - (-1/6) = 1/3.$$

Число $R(A, B)$ называется *коэффициентом регрессии* события A относительно события B и является удобной мерой зависимости A от B. Коэффициент регрессии $R(A, B)$ учитывает реализуемость явления, описываемого событием A, как при условии, что явление, описываемое событием B, реализуется, так и при условии, что это явление не реализуется.

Пример 2. В примере 2 из § 4 были вычислены вероятность $P(A)$ события A, описывающего появление среди выбранных наугад из урны шаров только красных, и условная вероятность $P_B(A)$ события A при условии B, описывающем появление шаров одного цвета:

$$P_B(A) = 1/6 \neq 1/30 = P(A).$$

Это неравенство описывает зависимость между появлением только красных шаров и появлением шаров одного цвета. Говорят также о зависимости события A от события B. Как и в примере 1, для измерения этой зависимости можно использовать коэффициент регрессии $R(A, B)$ события A относительно события B. Так как $AB' = 0$, то

$$P_{B'}(A) = n(AB')/n(B') = 0.$$

Поэтому

$$R(A, B) = P_B(A) - P_{B'}(A) = 1/6 - 0 = 1/6.$$

Сравним $R(A, B)$ с коэффициентом регрессии $R(B, A)$ события B относительно события A. Так как $BA = A$, то

$$P_A(B) = n(BA)/n(A) = n(A)/n(A) = 1.$$

Это равенство описывает тот факт, что все красные шары одного цвета.

Вычислим условную вероятность $P_{A'}(B)$ события B, описывающего появление шаров одного цвета, при условии A', описывающем появление не только красных шаров.

Пересечение BA' событий B и A' описывает появление только белых шаров. Используя проведенные раньше вычисления, получаем:

$$P_{A''}(B) = n(BA')/n(A') = 20/116 = 5/29.$$

Поэтому

$$R(B, A) = P_A(B) - P_{A'}(B) = 1 - 5/29 = 24/29.$$

Таким образом,

$$R(A, B) \neq R(B, A).$$

Коэффициент регрессии A относительно B не равен коэффициенту регрессии B относительно A.

Пример 3. В примере 3 из § 4 были вычислены вероятность $P(A)$ события $A = \{01, 11\}$, описывающего появление герба при втором подбрасывании монеты, и условная вероятность $P_B(A)$ события A при условии $B = \{10, 11\}$, описывающем появление герба при первом подбрасывании:

$$P_B(A) = 1/2 = P(A).$$

Аналогично,

$$P_{B'}(A) = n(AB')/n(B') = 1/2 = P(A).$$

Поэтому коэффициент регрессии события A относительно события B равен нулю:

$$R(A, B) = P_B(A) - P_{B'}(A) = 1/2 - 1/2 = 0.$$

В связи с этим говорят, что событие A не зависит от события B.

Равен нулю и коэффициент регрессии события B относительно события A. Действительно, так как $BA = \{11\}$, то

$$P_B(A) = n(BA)/n(A) = 1/2 = P(B).$$

Аналогично,

$$P_{A'}(B) = n(BA')/n(A') = 1/2 = P(B).$$

Поэтому

$$R(B, A) = P_A(B) - P_{A'}(B) = 1/2 - 1/2 = 0.$$

В связи с этим говорят, что событие B не зависит от события A.

В итоге можно сказать, что события A и B *независимы*. Это соответствует очевидной независимости результатов различных подбрасываний реальной монеты.

2.5. Определения

Примеры 1—3 подводят к следующим определениям зависимости событий в модели с множеством исходов U, алгеброй событий $\mathscr{A}$ и вероятностью P.

Рассмотрим произвольные *возможные* события A и B. Заметим, что равенства

$$P_B(A) = P(A), \quad P(AB) = P(A)P(B), \quad P_A(B) = P(B)$$

эквивалентны. В самом деле, если

$$P_B(A) = P(A),$$

то, по правилу умножения,

$$P(AB)=P(B)P_B(A)=P(B)P(A).$$

Обратно, если

$$P(AB)=P(A)P(B),$$

то

$$P_B(A)=P(AB)/P(B)=P(A)P(B)/P(B)=P(A).$$

Таким образом, равенства

$$P_B(A)=P(A), \quad P(AB)=P(A)P(B)$$

эквивалентны. Последнее из этих равенств симметрично относительно событий A и B. Поменяв их местами, убеждаемся в эквивалентности равенств

$$P_A(B)=P(B), \quad P(AB)=P(A)P(B).$$

Значит, все три рассматриваемые равенства эквивалентны: если верно одно из них, то верны и остальные. Для определения зависимости событий A и B удобно использовать среднее из этих равенств, симметричное относительно A и B.

Если верно равенство

$$P(AB)=P(A)P(B), \tag{9}$$

то говорят, что события A и B *независимы*. Если равенство (9) не верно, то говорят, что события A и B *зависимы*.

Это определение удобно распространить на случай, когда событие A или событие B невозможно. Тогда равенство (9) автоматически верно:

$$P(AB)=0=P(A)P(B).$$

Таким образом, если хотя бы одно из событий A и B невозможно, эти события независимы.

Примеры:

1) в примере 1 события A и B зависимы, так как

$$P(AB)=1/3\neq 1/2\cdot 1/2=P(A)P(B);$$

2) в примере 2 события A и B также зависимы, так как

$$P(AB)=1/6\neq 1/30\cdot 2/5=P(A)P(B);$$

3) в примере 3 события A и B независимы, так как

$$P(AB)=1/4=1/2\cdot 1/2=P(A)P(B).$$

Замечание. Данное определение зависимости событий не всегда согласуется с интуицией. В частности, в модели примера 1 с подбрасыванием игральной кости рассмотрим события $A_1=\{1, 2, 3\}$, $A_2=\{1, 2, 3, 4\}$ и $B=\{1, 3, 5\}$. Как было показано, события A_1 и B зависимы. В то же время события A_2 и B независимы:

$$P(A_2B)=1/3=2/3\cdot 1/2=P(A_2)P(B).$$

По-видимому, без вычислений трудно объяснить, почему появление меньше трех очков зависит от их нечетности, а появление меньше четырех — нет.

Если $A=B$, то равенство (9) превращается в равенство

$$P(A)=P(A)P(A).$$

Это равенство верно, если и только если верно одно из равенств

$$P(A)=0, \quad P(A)=1,$$

т. е. если A — невозможное или достоверное событие:

$$A=0, \quad A=U.$$

Таким образом, по данному определению, невозможное и достоверное события (и только они) не зависят от самих себя.

3.5. Коэффициенты регрессии

Для измерения зависимости событий определяются коэффициенты регрессии.

По-прежнему имеется в виду модель с множеством исходов U, алгеброй событий $\mathcal{A}$ и вероятностью P. Рассмотрим произвольные *возможные* события A и B, дополнительные события A' и B' для которых *тоже возможны*.

Хотя зависимость событий A и B определена симметрично, иногда важно учитывать различный характер зависимости A от B и B от A. В частности, в примере 2 с шарами появление шаров одного цвета (событие B) не обязательно означает появление только красных шаров (событие A): шары могут оказаться белыми. В то же время появление только красных шаров означает появление шаров одного цвета. Таким образом, в данном случае A меньше зависит от B, чем B от A. Предлагаемый в качестве меры зависимости событий коэффициент регрессии учитывает различный характер зависимости одного события от другого.

Заметим, что

$$P_B(A)-P_{B'}(A)=\frac{P(AB)-P(A)P(B)}{P(B)(1-P(B))}.$$

В самом деле, используя правило деления, получаем:

$$P_B(A)-P_{B'}(A)=\frac{P(AB)}{P(B)}-\frac{P(AB')}{P(B')}.$$

Как нетрудно проверить, используя определение вероятности P и правило вычитания для числа элементов,

$$P(B')=P(U-B)=P(U)-P(B)=1-P(B),$$

$$P(AB')=P(A(U-B))=P(AU-AB)=P(A-AB)=$$
$$=P(A)-P(AB).$$

Поэтому

$$P_B(A) - P_{B'}(A) = \frac{P(AB)}{P(B)} - \frac{P(A) - P(AB)}{1 - P(B)} = \frac{P(AB) - P(A)P(B)}{P(B)(1 - P(B))},$$

что и требовалось доказать.

Назовем *коэффициентом регрессии* события A относительно события B число

$$(10) \quad R(A, B) = P_B(A) - P_{B'}(A) = \frac{P(AB) - P(A)P(B)}{P(B)(1 - P(B))}.$$

Коэффициент регрессии $R(A, B)$ является *мерой зависимости* события A от события B. Он может быть не равен коэффициенту регрессии

$$(11) \quad R(B, A) = P_A(B) - P_{A'}(B) = \frac{P(AB) - P(A)P(B)}{P(A)(1 - P(A))},$$

являющемуся мерой зависимости события B от события A.

Примеры:

1) в примере 1 $R(A, B) = 1/3$, $R(B, A) = 1/3$;
2) в примере 2 $R(A, B) = 1/6$, $R(B, A) = 24/29$;
3) в примере 3 $R(A, B) = 0$, $R(B, A) = 0$.

4.5. Коэффициент корреляции

Используя коэффициенты регрессии, можно определить удобную симметричную меру зависимости событий. Этой мерой является *коэффициент корреляции.*

Будем рассматривать события A и B, удовлетворяющие условиям, сформулированным при определении коэффициентов регрессии.

Заметим, что

$$P(A) > 0, \quad 1 - P(A) = P(A') > 0, \quad P(B) > 0, \quad 1 - P(B) = P(B') > 0.$$

Поэтому из равенств (10) и (11) следует, что коэффициенты регрессии для событий A и B имеют одинаковые знаки:

$$R(A, B)R(B, A) \geqslant 0.$$

Следовательно, для них можно определить среднее геометрическое

$$\pm\sqrt{R(A, B)R(B, A)},$$

выбирая знак перед корнем равным знаку коэффициентов регрессии $R(A, B)$ и $R(B, A)$. Из равенств (10) и (11) вытекает, что

$$\pm\sqrt{R(A, B)R(B, A)} = \frac{P(AB) - P(A)P(B)}{\sqrt{P(A)(1 - P(A))P(B)(1 - P(B))}}.$$

Назовем *коэффициентом корреляции* событий A и B число

$$(12) \quad K(A, B) = \pm\sqrt{R(A, B)R(B, A)} =$$

$$= \frac{P(AB) - P(A)P(B)}{\sqrt{P(A)(1 - P(A))P(B)(1 - P(B))}}.$$

Ясно, что

$$K(A, B) = K(B, A).$$

Коэффициент корреляции событий A и B равен коэффициенту корреляции событий B и A.

Примеры:

1) в примере 1 $K(A, B) = 1/3$;
2) в примере 2 $K(A, B) = 2/29$;
3) в примере 3 $K(A, B) = 0$.

Сравнивая равенства (9), (10) и (11), видим, что коэффициенты регрессии для событий A и B равны нулю, если и только если эти события независимы.

Сравнивая равенства (9) и (12), видим, что коэффициент корреляции событий A и B равен нулю, если и только если они независимы.

Задачи, поясняющие использование модели Лапласа, собраны в главе 1 части III.

Глава 2

МОДЕЛЬ БЕРНУЛЛИ

Модель Бернулли описывает опыты, состоящие из последовательности нескольких одинаковых и независимых испытаний, каждое из которых имеет два возможных результата. Эта модель, как и модель Лапласа, является одной из простейших моделей теории вероятностей. Вместе с тем модель Бернулли достаточно богата в идейном отношении и имеет многочисленные приложения.

При первом чтении рекомендуется ограничиться параграфами 1 и 2.

§ 1. ПРИМЕРЫ

Рассмотрим несколько примеров, разъясняющих содержание модели Бернулли и подготавливающих соответствующие формальные определения.

В примерах главы 1 при описании реальных ситуаций использовались понятия *результат*, *явление* и *меры их реализуемости*. В предлагавшихся математических моделях эти понятия соответственно описывались терминами *исход*, *событие*, *элементарная вероятность* и *вероятность*. Те же понятия и термины будут использоваться в примерах главы 2.

1.1. Пример 1. Одно подбрасывание монеты

Рассмотрим испытание, заключающееся в подбрасывании симметричной монеты один раз, и построим вероятностную модель этого испытания.

1.1.1. Множество исходов $U=B$

Будем считать, что возможен один и только один из двух *результатов*: монета падает *гербом* вверх либо монета падает *цифрой* вверх. Условимся первый из этих результатов описывать числом 1, а второй — числом 0. Эти числа будем называть *исходами*. Составляемое этими числами множество исходов

$$U=B=\{1,\ 0\}$$

является удобным математическим описанием возможных результатов испытания с монетой.

2.1.1. Алгебра событий $\mathscr{A}=\mathscr{B}$

Явления, определенные результатами рассматриваемого испытания, условимся математически описывать множествами, составленными из соответствующих исходов.

Пустое множество O описывает *невозможное* событие: не появляется герб и не появляется цифра. Элементарное множество $У=\{1\}$ описывает появление герба. Элементарное множество $H=\{0\}$ — появление цифры. Множество исходов U описывает *достоверное* событие: либо появляется герб, либо появляется цифра. Употребление букв $У$ и H связано с тем, что при игре с монетой обычно появление герба означает выигрыш (*успех*), а появление цифры — проигрыш (*неудачу*).

Логические связи между явлениями описываются объединением, дополнением и пересечением соответствующих множеств исходов.

В частности, невозможность одновременного появления герба и цифры описывается равенством

$$УH=O,$$

достоверность появления герба либо цифры — равенством

$$У+H=U.$$

То, что непоявление герба означает появление цифры и непоявление цифры означает появление герба, описывается равенствами

$$У'=U-У=H, \quad H'=U-H=У.$$

Каждое множество, составленное из исходов, условимся называть *событием*. Класс

$$\mathscr{A}=\mathscr{B}=\{O,\ У,\ H,\ U\}$$

всех таких множеств вместе с их объединением, пересечением и дополнением образует *алгебру* событий $(\mathscr{A}, \cup, \cap, ')$. Она служит удобным математическим описанием явлений, определяемых возможными результатами испытания с монетой.

3.1.1. Элементарная вероятность $p \equiv q$

Симметричность монеты позволяет считать результаты рассматриваемого испытания *равновозможными*.

Поэтому появлению герба, которое зависит от случая, в качестве *меры реализуемости* можно приписать число $a=1/2$, а появлению цифры — также число $1-a=1/2$. В соответствии с этим определим элементарную вероятность $p(1)=q(1)$ исхода 1 равенством

$$p(1)=q(1)=a=1/2,$$

а элементарную вероятность $p(0)=q(0)$ исхода 0 — равенством

$$p(0)=q(0)=1-a=1/2.$$

Определенная этими равенствами функция $p=q$ на множестве исходов $U=B$ называется ***элементарной вероятностью*** и служит удобным математическим описанием меры реализуемости возможных результатов испытания с монетой.

4.1.1. Вероятность $P=Q$

Каждому явлению, определяемому возможными результатами рассматриваемого испытания, естественно в качестве *меры реализуемости* приписать сумму мер тех результатов испытания, при которых это явление реализуется. В соответствии с этим определим вероятности событий O, $У$, H, U равенствами

$$P(O)=Q(O)=0, \quad P(У)=Q(У)=a=1/2,$$

$$P(H)=Q(H)=1-a, \quad P(U)=Q(U)=1.$$

Определенная этими равенствами функция $P=Q$ на классе событий $\mathscr{A}=\mathscr{B}$ называется ***вероятностью*** и служит удобной мерой реализуемости явлений, определяемых возможными результатами испытания с монетой.

5.1.1. Вероятностная модель $(B, \mathscr{B};\ q,\ Q)$

Множество исходов $U=B$, алгебра событий $\mathscr{A}=\mathscr{B}$, элементарная вероятность $p=q$ и вероятность $P=Q$ образуют удобную модель испытания с подбрасыванием симметричной монеты один раз. Это — *модель Бернулли* $n=1$ испытания с ***вероятностью успеха*** $a=1/2$. Она совпадает с классической моделью Лапласа для $2^n=2$ равновероятных исходов.

2.1. Пример 2. Один выстрел по цели

Рассмотрим испытание, заключающееся в выстреле по цели один раз, и построим вероятностную модель этого испытания.

1.2.1. Множество исходов $U=B$

Будем считать, что возможен один и только один из *результатов*: *попадание* либо *промах*. Все другие результаты (например, осечка) исключаются. В этом случае возможные результаты рассматриваемого испытания, как и в примере 1, хорошо описываются множеством исходов $U=B=\{1, 0\}$. Исход 1 описывает попадание, а исход 0 — промах.

2.2.1. Алгебра событий $\mathscr{A}=\mathscr{B}$

Явления, определяемые возможными результатами выстрела, как и в примере 1, хорошо описываются ***алгеброй*** событий $\mathscr{A}=\mathscr{B}$, образованной ***событиями*** O, $У=\{1\}$, $H=\{0\}$, $U=B$ и их объединением, пересечением и дополнением. События $У$ и H описывают соответственно попадание и промах. Употребление букв $У$ и H связано с представлением о попадании как об успехе, а о промахе — как неудаче.

3.2.1. Элементарная вероятность $p=q$

Как правило, попадание и промах при выстреле нельзя считать равновозможными, как появление герба и цифры при подбрасывании монеты. Поэтому элементарную вероятность, являющуюся мерой реализуемости попадания и промаха, нужно определить иначе, чем в примере 1 для герба и цифры.

Предположим, что в таких же условиях прежде было произведено большое число n выстрелов, m из которых привели к попаданию, а $n-m$ — к промаху. Естественно ожидать, что одинаковость условий и в дальнейшем обеспечит частоту попадания, близкую наблюдавшейся прежде частоте $a=m/n$, и использовать эту частоту в качестве ***меры реализуемости*** для попадания, а возможность промаха оценивать частотой промаха $1-a$.

В частности, если $n=1000$ и $m=900$, то $a=9/10$ и $1-a=1/10$.

В соответствии с этим определим вероятность $p(1)$ исхода 1 равенства

$$p(1)=q(1)=a,$$

а элементарную вероятность $p(0)$ исхода 0 — равенством

$$p(0)=q(0)=1-a.$$

Определенная этими равенствами функция $p=q$ на множестве исходов $U=B$ называется ***элементарной вероятностью*** и служит удобным математическим описанием меры реализуемости для попадания и промаха.

4.2.1. Вероятность $P=Q$

Используя элементарные вероятности исходов, определим вероятности событий O, $У$, H, U равенствами

$$P(O)=Q(O)=0, \quad P(У)=Q(У)=a,$$
$$P(H)=Q(H)=1-a, \quad P(U)=Q(U)=1.$$

Определяемая этими равенствами функция $P=Q$ на алгебре событий $\mathscr{A}=\mathscr{B}$ называется *вероятностью* и служит удобным математическим описанием меры реализуемости явлений, определяемых возможными результатами выстрела.

5.2.1. Вероятностная модель $(B, \mathscr{B}; q, Q)$

Множество исходов $U=B$, алгебра событий $\mathscr{A}=\mathscr{B}$, элементарная вероятность $p=q$ и вероятность $P=Q$ образуют удобную модель выстрела по цели. Это — *модель Бернулли* $n=1$ испытания с *вероятностью успеха* $a=9/10$. Она не совпадает с классической моделью Лапласа для $2^n=2$ равновероятных исходов.

3.1. Пример 3. Два подбрасывания монеты

Рассмотрим опыт, состоящий из *последовательности* двух *испытаний*, каждое из которых заключается в подбрасывании симметричной монеты один раз и имеет два возможных *результата*: появление *герба* либо появление *цифры*. Построим вероятностную модель этого опыта.

Опыт с подбрасыванием два раза симметричной монеты рассматривался в примере 3 из § 1 главы 1. Там было показано, что этот опыт хорошо описывается классической моделью Лапласа для четырех равновозможных исходов. К той же самой модели можно прийти, компонуя две модели примера 1, описывающие каждая подбрасывание один раз симметричной монеты.

1.3.1. Множество исходов $U=B\times B$

Как показано в примере 1, результаты подбрасывания один раз монеты можно описать множеством

$$B=\{1, 0\}.$$

Значит, результаты подбрасывания два раза монеты можно описать множеством *исходов*

$$U=B\times B=\{11, 10, 01, 00\}.$$

Это то же самое множество, что и в примере 3 из §1 главы 1.

Для большей выразительности условимся появление герба считать *успехом*, а появление цифры — *неудачей* и говорить:

успех при первом испытании и *успех* при втором испытании;

успех при первом испытании и *неудача* при втором испытании;

неудача при первом испытании и *успех* при втором испытании;

неудача при первом испытании и *неудача* при втором испытании;

вместо, соответственно,

герб при первом подбрасывании и *герб* при втором подбрасывании;

герб при первом подбрасывании и *цифра* при втором подбрасывании;

цифра при первом подбрасывании и *герб* при втором подбрасывании;

цифра при первом подбрасывании и *цифра* при втором подбрасывании.

Математическое описание возможных результатов рассматриваемого опыта выражает следующая

Таблица исходов

Успех при первом испытании и *успех* при втором испытании	1	1
Успех при первом испытании и *неудача* при втором испытании	1	0
Неудача при первом испытании и *успех* при втором испытании	0	1
Неудача при первом испытании и *неудача* при втором испытании	0	0

2.3.1. Алгебра событий $\mathscr{A}=\mathscr{B}\times\mathscr{B}$

Как показано в примере 1, *явления*, определяемые результатами подбрасывания один раз монеты, можно описать алгеброй множеств класса

$$\mathscr{B}=\{O,\ У,\ Н,\ В\},$$

порожденного элементарными множествами $У=\{1\}$ и $Н=\{0\}$. Все остальные множества $\mathscr{B}$ получаются из множеств $У$ и $Н$ с помощью алгебраических действий:

$$УН=O, \quad У+Н=В.$$

Значит, явления, определяемые результатами подбрасывания два раза монеты, можно описать *алгеброй событий* $\mathscr{A}$, порожденной элементарными множествами

$$У\times У=\{11\}, \quad У\times Н=\{10\}, \quad Н\times У=\{01\}, \quad Н\times Н=\{00\}.$$

В связи с этим условимся обозначать алгебру событий $\mathscr{A}$ также символом $\mathscr{B}\times\mathscr{B}$ и называть ее *произведением* алгебр $\mathscr{B}$. Это та же алгебра, что и в примере 3 из § 1 главы 1, так как каждая часть A множества $U=B\times B$ является объединением некоторых из множеств $У\times У$, $У\times Н$, $Н\times У$, $Н\times Н$.

Таблица исходов показывает, что события $У\times У$, $У\times Н$, $Н\times У$, $Н\times Н$, как и составляющие их исходы 11, 10, 01, 00, описывают результаты рассматриваемого опыта.

Для описания каждого подбрасывания монеты удобно ввести следующие обозначения:

Таблица событий

Успех при первом испытании	$У_1 = У \times B = \{11, 10\}$
Неудача при первом испытании	$H_1 = H \times B = \{01, 00\}$
Успех при втором испытании	$У_2 = B \times У = \{11, 01\}$
Неудача при втором испытании	$H_2 = B \times H = \{10, 00\}$

Явления, определенные результатами первого подбрасывания монеты, описываются классом

$$\mathscr{B}_1 = \{O, У_1, H_1, U\},$$

который порождается множествами $У_1$ и H_1:

$$У_1 H_1 = O, \quad У_1 + H_1 = U.$$

Аналогично явления, определяемые результаты второго подбрасывания монеты, описываются классом

$$\mathscr{B}_2 = \{O, У_2, H_2, U\},$$

который порождается множествами $У_2$ и H_2:

$$У_2 H_2 = O, \quad У_2 + H_2 = U.$$

Заметим, что

$$У_1 У_1 = У \times У = \{11\}, \quad У_1 H_2 = У \times H = \{10\},$$
$$H_1 У_2 = H \times У = \{01\}, \quad H_1 H_2 = H \times H = \{00\}.$$

Поэтому все события получаются из событий $У_1$, H_1, $У_2$, H_2 с помощью алгебраических операций, т. е. алгебра $\mathscr{A}$ порождается алгебрами $\mathscr{B}_1$ и $\mathscr{B}_2$. Это соответствует тому, что рассматриваемый опыт состоит из первого и второго испытаний. Конкретные выражения событий дает следующая

Таблица событий

$У_1 H_2 + H_1 H_2 = \{10, 01\}$	$У_1 + H_1 H_2 = \{11, 10, 00\}$
$У_1 У_2 + H_1 H_2 = \{11, 00\}$	$H_1 + У_1 У_2 = \{01, 00, 11\}$
$У_1 + H_1 У_2 = \{11, 10, 01\}$	$H_1 + У_1 H_2 = \{01, 00, 10\}$

Невозможное событие O, достоверное событие U и события, фигурирующие в приведенных таблицах, исчерпывают *алгебру событий* $\mathscr{A}$.

Упражнение. Дать словесные описания явлений, соответствующих событиям из последней таблицы.

3.3.1. Элементарная вероятность $p = q \times q$

Заметим, что шансы на появление герба при первом и втором подбрасываниях равны. Точно так же одинаково возможно появление цифры при первом и втором подбрасываниях. Другими словами, равны шансы на успех при первом и втором испытаниях, а также шансы на неудачу при первом и втором испытаниях. В этом смысле испытания, составляющие рассматриваемый опыт, можно назвать *одинаковыми*. Соответственно в описывающей опыт модели должны быть равны вероятности успеха при первом и втором испытаниях и вероятности неуспеха при первом и втором испытаниях:

$$(1') \qquad P(У_1) = P(У_2), \quad P(H_1) = P(H_2).$$

Заметим также, что шансы на появление герба при одном подбрасывании не зависят от результата другого. Точно также возможность появления цифры при одном подбрасывании не зависит от результата другого. Другими словами, шансы на успех при одном испытании не зависят от результата другого. Также возможность неудачи при одном испытании не зависит от результата другого. В этом смысле испытания, составляющие рассматриваемый опыт, можно назвать *независимыми*. Соответственно в описывающей опыт модели события $У_1$ и $У_2$, $У_1$ и H_2, H_1 и $У_2$, H_1 и H_2 должны быть независимы. По аналогии с моделью Лапласа будем выражать независимость событий с помощью правила умножения:

$$(2') \qquad P(У_1У_2) = P(У_1)P(У_2), \quad P(У_1H_2) = P(У_1)P(H_2),$$
$$P(H_1У_2) = P(H_1)P(У_2), \quad P(H_1H_2) = P(H_1)P(H_2).$$

Используя равенства (1′), (2′) и модель примера 1, можно определить элементарную вероятность p и вероятность P для создаваемой модели.

Как показано в примере 1, вероятность успеха и вероятность неудачи при испытании, заключающемся в подбрасывании монеты один раз, выражаются соответственно равенствами

$$Q(У) = q(1) = a, \quad Q(H) = q(0) = 1 - a,$$

где $a = 1/2$. Следовательно, в силу равенств (1′)

$$P(У_1) = P(У_2) = Q(У) = q(1) = a,$$
$$P(H_1) = P(H_2) = Q(H) = q(0) = 1 - a.$$

События $У_1У_2 = \{11\}$, $У_1H_2 = \{10\}$, $H_1У_2 = \{01\}$, $H_1H_2 = \{00\}$ являются элементарными. Поэтому из полученных равенств и равенств (2′) следует, что

$$(3') \qquad \begin{array}{ll} p(11) = q(1)q(1) = aa, & p(10) = q(1)q(0) = a(1-a), \\ p(01) = q(0)q(1) = (1-a)a, & p(00) = q(0)q(0) = (1-a)(1-a). \end{array}$$

Равенства (3′) определяют функцию $p=q\times q$ на множестве исходов $U=B\times B$, которая является *элементарной вероятностью* для создаваемой модели и мерой реализуемости возможных результатов рассматриваемого опыта. Это та же элементарная вероятность, что и в примере 1 из § 1 главы 1.

4.3.1. Вероятность $P=Q\times Q$

Вероятность P для создаваемой модели, как и для классической модели Лапласа, естественно определить с помощью элементарной вероятности p и правила сложения. Для каждого события A вероятность $P(A)$ события A определяется как сумма элементарных вероятностей $p(u)$ исходов u, составляющих событие A:

$$P(A)=\sum_{u\in A} p(u). \tag{4′}$$

В частности, в соответствии с равенствами (1) верны равенства

$$P(У_1)=P(У_2)=a, \quad P(H_1)=P(H_2)=1-a \tag{1″}$$

и в соответствии с равенствами (2) — равенства

$$\begin{gathered} P(У_1У_2)=P(У_1)P(У_2)=aa, \\ P(У_1H_2)=P(У_1)P(H_2)=a(1-a), \\ P(H_1У_2)=P(H_1)P(У_2)=(1-a)a, \\ P(H_1H_2)=P(H_1)P(H_2)=(1-a)(1-a). \end{gathered} \tag{2″}$$

Таким образом, предлагаемая модель учитывает одинаковость и независимость испытаний, составляющих рассматриваемый опыт.

Используя равенства (1″), (2″) и правило сложения, легко вычислить вероятность событий, составляющих таблицу событий:

Таблица вероятностей

$P(У_1H_2+H_1У_2)=a(1-a)+(1-a)a$	$P(У_1+H_1H_2)=a+(1-a)(1-a)$
$P(У_1У_2+H_1H_2)=aa+(1-a)(1-a)$	$P(H_1+У_1У_2)=(1-a)+aa$
$P(У_1+H_1У_2)=a+(1-a)a$	$P(H_1+У_1H_2)=(1-a)+a(1-a)$

Так как $a=(1-a)=1/2$, то в первом столбце таблицы находятся числа 1/2, 1/2, 3/4, а во втором — 3/4, 3/4, 3/4.

Для того, чтобы подчеркнуть связь между вероятностью Q для одного испытания и вероятностью P для двух одинаковых и независимых испытаний, выражаемую равенствами (2′), условимся обозначать вероятность P также символом $Q\times Q$ и называть ее *произведением* вероятностей Q.

5.3.1. Вероятностная модель $(B\times B,\ \mathscr{B}\times\mathscr{B},\ q\times q,\ Q\times Q)$

Множество исходов $U=B\times B$, алгебра событий $\mathscr{A}=\mathscr{B}\times\mathscr{B}$, элементарная вероятность $p=q\times q$ и вероятность $P=Q\times Q$ обра-

зуют удобную модель подбрасывания два раза симметричной монеты.

Эта модель является моделью последовательности двух одинаковых и независимых испытаний, каждое из которых имеет два равновозможных результата. Условимся называть ее *моделью Бернулли* $n=2$ испытаний с *вероятностью успеха* $a=1/2$. Она совпадает с классической моделью Лапласа для $2^n=4$ равновероятных исходов.

4.1. Пример 4. Два выстрела по цели

Рассмотрим опыт, состоящий из *последовательности* двух *испытаний*, каждое из которых заключается в выстреле по цели один раз и имеет два возможных *результата*: *попадание* либо *промах*. Построим вероятностную модель этого опыта.

По аналогии с примером 3 будем строить модель для двух выстрелов, компонуя две модели примера 2, описывающие каждая один выстрел.

1.4.1. Множество исходов $U=B\times B$

Так же, как в примере 3 с двумя подбрасываниями монеты, возможные *результаты* опыта с двумя выстрелами описывают множество *исходов* $U=B\times B$. По аналогии с примером 3 условимся считать попадание *успехом*, а промах — *неудачей*. Тогда предлагаемое описание выражается таблицей исходов примера 3.

2.4.1. Алгебра событий $\mathcal{A}=\mathcal{B}\times\mathcal{B}$

Так же, как в примере 3, *явления*, определяемые возможными результатами рассматриваемого опыта, описывает алгебра *событий* $\mathcal{A}=\mathcal{B}\times\mathcal{B}$, образованная всеми частями множества исходов U. Это описание на языке успехов и неудач выражают таблицы исходов и событий.

3.4.1. Элементарная вероятность $p=q\times q$

Предположим, что выстрелы производятся в одинаковых условиях. Возможности попадания при первом и втором выстреле равны, так же, как и возможности промаха. Другими словами, предположим, что составляющие рассматриваемый опыт испытания *одинаковы*.

Предположим, что шансы на попадание так же, как и возможность промаха при одном выстреле, не зависят от результата другого. Т. е. предположим, что составляющие рассматриваемый опыт испытания *независимы*.

Предположим, наконец, что известна *частота* a попадания при достаточно большом числе выстрелов, произведенных в таких же условиях.

Рассуждая по аналогии с примером 3 и используя равенства (1′), (2′) и модель примера 2 для рассматриваемого опыта, определим *элементарную вероятность* $p=q\times q$ равенствами (3′).

4.4.1. Вероятность $P=Q\times Q$

Вероятность $P=Q\times Q$ определяется равенством (4′).

Для предлагаемой модели, как и для модели примера 3, верны равенства (1″) и (2″). Следовательно, как и модель примера 3, она учитывает одинаковость и независимость испытаний, составляющих рассматриваемый опыт.

5.4.1. Вероятностная модель ($B\times B$, $\mathscr{B}\times\mathscr{B}$; $q\times q$, $Q\times Q$)

Множество исходов $U=B\times B$, алгебра событий $\mathscr{A}=\mathscr{B}\times\mathscr{B}$ и вероятность $P=Q\times Q$ образуют удобную модель двух выстрелов по цели. Эта модель является моделью последовательности двух одинаковых и независимых испытаний, каждое из которых имеет два возможных результата. Условимся называть ее *моделью Бернулли* $n=2$ испытаний с *вероятностью успеха* a ($0\leqslant a\leqslant 1$). Если $a\neq 1/2$, то эта модель не совпадает с классической моделью Лапласа для $2^n=4$ равновероятных исходов.

Например, если $a=9/10$, то равенства (3′) для элементарных вероятностей исходов дают следующие значения:

$$p(11)=81/100, \quad p(10)=p(01)=9/100, \quad p(00)=1/100.$$

§ 2. МОДЕЛЬ БЕРНУЛЛИ

Для описания опытов, составленных из *последовательности* нескольких *одинаковых* и *независимых испытаний*, каждое из которых имеет *два возможных результата* (успех либо неудача), удобна модель Бернулли.

1. 2. Множество исходов $U=B^n$

Рассмотрим произвольное натуральное число $n>0$ и множество

$$B=\{1, 0\}.$$

Основой модели Бернулли является множество

$$U=B^n$$

всех строк $u=(u_j)_{1\leqslant j\leqslant n}$ длины n, составленных из элементов 1 и 0 множества B. Эти строки называются *исходами*, а множество U — *множеством исходов*. Множество исходов U описывает возможные *результаты* рассматриваемого опыта.

Строка

$$u=(u_j)_{1\leqslant j\leqslant n}=u_1 \ \dots \ u_n,$$

на j-м месте которой расположен элемент u_j, равный 1 или 0, описывает результат опыта, при котором результатом j-го испытания является *успех*, если $u_j=1$, и *неудача*, если $u_j=0$.

Например, строка

$$u=1\dots1$$

описывает результат опыта, при котором результатом каждого испытания является успех, а строка

$$u=0\dots0$$

— результат опыта, при котором результатом каждого испытания является неудача. Полное описание результатов опыта, составленного из $n=2$ испытаний, дается таблицей исходов примера 3 из § 1.

2.2. Алгебра событий $\mathscr{A}=\mathscr{B}^n$

Вторым элементом модели Бернулли является алгебра $\mathscr{A}=\mathscr{B}^n$, образованная классом A всех частей множества исходов U и определенными для этих частей объединением, пересечением и дополнением. Части A множества исходов U называются *событиями*. Алгебра $(\mathscr{A}, \bigcup, \bigcap, ')$ называется алгеброй событий и сокращенно обозначается буквой $\mathscr{A}$ (так же, как и класс событий). Алгебра событий $(\mathscr{A}, \bigcup, \bigcap, ')$ описывает *явления*, определяемые возможными результатами рассматриваемого опыта.

Рассмотрим элементарные множества

$$У=\{1\},\quad H=\{0\},$$

класс множеств

$$\mathscr{B}=\{O,\ У,\ H,\ B\}$$

и образованную этим классом, объединением, пересечением и дополнением алгебру $(\mathscr{B}\ \bigcup,\ \bigcap, ')$. Сокращенно эту алгебру условимся обозначать символом $\mathscr{B}$ (так же, как и множество ее элементов). Алгебра $\mathscr{B}$ порождается множествами $У$ и H:

$$УH=O,\quad У+H=B.$$

Специфика алгебры событий $\mathscr{A}$ для модели Бернулли заключается в том, что она является произведением n алгебр $\mathscr{B}$. Это значит, что каждое событие класса $\mathscr{A}$ может быть получено с помощью алгебраических операций из некоторых декартовых произведений n множеств класса $\mathscr{B}$. Действительно, для каждого события A класса верно равенство

$$A=\sum_{u\in A}\ \prod_{1\leqslant j\leqslant n}\{u_j\}.$$

В частности, для каждого элементарного события $\{u\}$ верно равенство

$$\{u\}=\prod_{1\leqslant j\leqslant n}\{u_j\}.$$

В связи с декартовым произведением удобно использовать геометрический язык и называть декартово произведение

$$X = \prod_{1 \leqslant j \leqslant n} X_j$$

n множеств X_j класса B_j *прямоугольником* со сторонами X_j. В частности, каждое элементарное событие $\{u\}$ является прямоугольником со сторонами $\{u_j\}$. Поэтому каждое событие A равно сумме некоторого семейства попарно не пересекающихся прямоугольников.

Среди различных прямоугольников выделяются прямоугольники

$$У_i = \{u : u_i = 1\}, \quad H_i = \{u : u_i = 0\} \quad (i = 1, \ldots, n),$$

составленные из всех строк u, у которых на i-м месте находятся соответственно 1 и 0. События $У_i$ и H_i описывают соответственно *успех* и *неудачу при i-м испытании*. Эти события являются прямоугольниками, i-е стороны которых равны соответственно $\{1\}$ и $\{0\}$, а остальные — B:

$$Y_i = \prod_{1 \leqslant j \leqslant n} X_{ij}, \quad X_{ii} = \{1\}, \quad X_{ij} = B \quad (i \neq j);$$

$$H_i = \prod_{1 \leqslant j \leqslant n} Y_{ij}, \quad Y_{ii} = \{0\}, \quad Y_{ij} = B \quad (i \neq j).$$

Ясно, что

$$У_i H_i = O, \quad У_i + H_i = U \quad (i = 1, \ldots, n).$$

Условимся пустое множество O называть *невозможным* событием, а множество исходов U — *достоверным* событием. Последние равенства означают, что в модели Бернулли успех *и* неудача при одном и том же испытании невозможны, а успех *или* неудача достоверны.

Полное описание событий $У_i$ и H_i для модели Бернулли $n=2$ испытаний содержится в таблице событий примера 3 из § 1.

Для каждого номера $i = 1, \ldots, n$ рассмотрим класс множеств

$$\mathscr{B}_i = \{O, У_i, H_i, U\},$$

порождаемый событиями $У_i$ и H_i. Классы $\mathscr{B}_i$ порождают алгебру событий $\mathscr{A}$: каждое событие A может быть получено с помощью алгебраических операций из событий $У_i$ и H_i $(i = 1, \ldots, n)$. Это соответствует тому, что рассматриваемый опыт состоит из $1, \ldots, n$-го испытаний. Условимся класс $\mathscr{B}_i$ называть *i-м испытанием*, а события $У_i$ и H_i — *исходами i-го испытания.*

Примеры конкретного выражения событий через исходы испытаний для модели Бернулли $n=2$ испытаний содержатся в таблице событий примера 3 из § 1.

3.2. Элементарная вероятность $p = q^n$

Третьим элементом модели Бернулли является функция $p = q^n$ на множестве исходов U, определяемая следующим образом.

Рассмотрим произвольное положительное число a, меньшее единицы:

$$0 \leqslant a \leqslant 1.$$

Определим значения $q(1)$, $q(0)$ функции q на множестве $B = \{1, 0\}$ равенствами

(1) $$q(1) = a, \quad q(0) = 1 - a.$$

Функция q на множестве B определяет функцию $p = q^n$ на множестве исходов $U = B^n$, значение $p(u)$ которой для каждого исхода $u = (u_j)_{1 \leqslant j \leqslant n}$ определяется равенством

(2) $$p(u) = \prod_{1 \leqslant j \leqslant n} q(u_j).$$

В частности, при $n = 2$ из равенств (1) и (2) следуют равенства (3′) примера 3 из § 1.

Функция p называется *элементарной вероятностью*, а ее значение $p(u)$ — элементарной вероятностью исхода u. Элементарная вероятность описывает зависимость возможных результатов опыта от случая. Она является *мерой реализуемости* этих результатов. Равенства (1) и (2) описывают *одинаковость* и *независимость* составляющих опыт испытаний.

Заметим, что из равенств (1) и (2) следует равенство

(3) $$p(u) = a^{s(u)} (1 - a)^{n - s(u)},$$

где

$$s(u) = \sum_{1 \leqslant j \leqslant n} u_j.$$

В частности, при $n = 2$ равенство (3) эквивалентно равенствам (3) примера 3 из § 1.

4.2. Вероятность $P = Q^n$

Четвертым и самым важным элементом модели Бернулли является функция

$$P = Q^n$$

на классе событий $\mathscr{A}$, определяемая следующим образом.

Для каждого события A рассмотрим сумму $P(A)$ элементарных вероятностей $p(u)$ исходов u, составляющих событие A:

(4) $$P(A) = \sum_{u \in A} p(u).$$

Равенство (4) определяет на классе событий $\mathscr{A}$ функцию P со значением $P(A)$ для каждого события A. Функция P называется *вероятностью*, а ее значение $P(A)$ — *вероятностью события* A. Вероятность описывает зависимость от случая явлений, определяемых результатами опыта. Она является *мерой реализуемости* этих явлений.

1.4.2*. Правило прямоугольника

Рассмотрим вероятность Q на классе множеств $\mathscr{B}$:

$$Q(O)=0,\quad Q(У)=a,\quad Q(H)=1-a,\quad Q(B)=1.$$

Специфика вероятности P для модели Бернулли заключается в том, что она является произведением n вероятностей Q. Это значит, что, аналогично тому, как в геометрии площадь прямоугольника равна произведению длин его сторон, в модели Бернулли вероятность $P(X)$ каждого прямоугольника X равна произведению вероятностей $Q(X_j)$ его сторон X_j $(j=1,\ldots,n)$, т. е. верно следующее

Правило прямоугольника:

$$P\left(\prod_{1\leqslant j\leqslant n} X_j\right)=\prod_{1\leqslant j\leqslant n} Q(X_j).$$

Доказательство. Используем принцип индукции. Рассмотрим множество M всех натуральных чисел m таких, что для $n=m+1$ верно правило прямоугольника.

1. Число $0\in M$, так как для $n=1$ правило прямоугольника эквивалентно равенству $P=Q^1=Q$;

2. Для каждого номера l верно предложение: если $l\in M$, то $l+1\in M$. Докажем это. Рассмотрим произвольный прямоугольник Y со сторонами $X_1,\ldots,X_n, X_{n+1}$ $(n=l+1)$. Обозначим прямоугольник со сторонами $X_1,\ldots,X_n$ буквой X:

$$X=\prod_{1\leqslant j\leqslant n} X_j,\quad Y=X\times X_{n+1}.$$

Рассмотрим произвольную строку $v\in Y$ и ее часть $u\in X$:

$$u=u_1\ldots u_n,\quad v=u_1\ldots u_n u_{n+1}=uu_{n+1}.$$

Из равенств (2) и (4) для $p=q^{n+1}$ и $P=Q^{n+1}$ вытекает, что

$$Q^{n+1}(Y)=\sum_{v\in Y} q^{n+1}(v)=\sum_{u\in X,\ u_{n+1}\in X_{n+1}} q^n(u)\,q(u_{n+1}).$$

Заметим, что

$$\sum_{u\in X,\ u_{n+1}\in X_{n+1}} q^n(u)\,q(u_{n+1})=\sum_{u\in X} q^n(u)\sum_{u_{n+1}\in X_{n+1}} q(u_{n+1}).$$

В самом деле, возможны четыре случая:

$$X_{n+1}=O,\quad X_{n+1}=\{1\},\quad X_{n+1}=\{0\},\quad X_{n+1}=B.$$

В первом случае обе части равны нулю, так как содержат суммы пустых семейств. Во втором и третьем случаях дело сводится к вынесению в левой части общих множителей $q(1)$ и $q(0)$. Наконец, в четвертом случае верны равенства

$$\sum_{u\in X,\ u_{n+1}\in X_{n+1}} q^n(u)\,q(u_{n+1})=\sum_{u\in X} q^n(u)\,q(1)+\sum_{u\in X} q^n(u)\,q(0)=$$

$$=\sum_{u\in X} q^n(u)\,(q(1)+q(0))=\sum_{u\in X} q^n(u)\sum_{u_{n+1}\in X_{n+1}} q(u_{n+1}).$$

Из равенств (2) и (4) для $p=q^n$ и $P=Q^n$, $p=q$ и $P=Q$ **следует**, что

$$\sum_{u\in X} q^n(u) \sum_{u_{n+1}\in X_{n+1}} q(u_{n+1}) = Q^n(X)\, Q(X_{n+1}).$$

Если $l\in M$, то для $n=l+1$ верно равенство

$$Q^n(X) = \prod_{1\leqslant j\leqslant n} Q(X_j).$$

Из полученных равенств вытекает, что в этом случае

$$Q^{n+1}(Y) = \prod_{1\leqslant j\leqslant n} Q(X_j)\, Q(X_{n+1}) = \prod_{1\leqslant j\leqslant n+1} Q(X_j)$$

и, значит, $l+1\in M$.

По принципу индукции из пунктов 1 и 2 следует, что $M=N$, т. е. правило прямоугольника верно для каждого номера $n\geqslant 1$.

Пример. Пусть $n=2$ и стороны прямоугольника Y определяются равенствами

$$X_1=\{1\},\quad X_2=\{1, 0\},\quad X_3=\{1, 0\}.$$

В этом случае

$$X=\{1\}\times\{1, 0\}=\{11, 10\},$$

$$Y=\{1\}\times\{1, 0\}\times\{1, 0\}=\{11, 10\}\times\{1, 0\}=\{111, 101, 110, 100\}.$$

$$u=u_2,\quad v=u_1u_2u_3=uu_3\ \ (u_1=1,\ u_2,\ u_3\in\{1, 0\}),$$

$$Q^3(Y)=q^3(111)+q^3(101)+q^3(110)+q^3(100)=$$

$$=q^2(11)q(1)+q^2(10)q(1)+q^2(11)q(0)+q^2(10)q(0)=$$

$$=[q^2(11)+q^2(10)][q(1)+q(0)]=Q^2(X)Q(X_3),$$

$$Q^2(X)=q^2(11)+q^2(10)=q(1)q(1)+q(1)q(0)=$$

$$=q(1)(q(1)+q(0))=Q(X_1)Q(X_2).$$

Следовательно,

$$Q^3(Y)=Q(X_1)Q(X_2)Q(X_3).$$

2.4.2*. Формула успехов и неудач

Из правила прямоугольника вытекает простое равенство для вероятности пересечения произвольных исходов каждых испытаний.

Рассмотрим произвольные *непересекающиеся* части K и L множества $\{1, \ldots, n\}$, составленные соответственно из k и l номеров. Множества K и L определяют событие

$$\bigcap_{i\in K} У_i \cdot \bigcap_{j\in L} H_j,$$

описывающее появление *успеха* при каждом испытании с номером $i\in K$ и *неудачи* при каждом испытании с номером $j\in L$.

Это событие является прямоугольником

$$\bigcap_{i\in K} У_i \cdot \bigcap_{j\in L} H_j = \prod_{1\leqslant m\leqslant n} X_m,$$

где

$$X_i = \{1\} \ (i \in K), \quad X_j = \{0\} \ (j \in L), \quad X_m = B \ (m \notin K+L).$$

Пример. Если $\{1, \ldots, n\} = \{1, 2, 3, 4, 5, 6\}$, $K = \{2, 4\}$, $L = \{6\}$ $(k=2,\ l=1)$,
то

$$\bigcap_{i \in K} У_i \bigcap_{j \in L} H_j = У_2 У_4 H_6 = B \times \{1\} \times B \times \{1\} \times B \times \{0\}.$$

Это событие описывает появление успеха при втором и четвертом испытаниях и неудачи при шестом испытании. Исходы остальных (первого, третьего, пятого) испытаний произвольны.

Формула успехов и неудач:

$$P\left(\bigcap_{i \in K} У_i \bigcap_{j \in L} H_j\right) = a^k (1-a)^l.$$

Доказательство. Из определения событий $У_i$, H_j и правила прямоугольника следует, что

$$P\left(\bigcap_{i \in K} У_i \bigcap H_j\right) = Q(\{1\})^k Q(\{0\})^l Q(B)^{n-k-l} = a^k (1-a)^l.$$

Формула успехов и неудач доказана.

Пример. $P(У_2 У_4 H_6) = a^2(1-a)$.

Замечание. Предположение о том, что множества K и L не пересекаются, существенно. Если K и L пересекаются, то в левой части формулы находится вероятность невозможного события O (успех и неудача при одном и том же испытании) и при $0 < a < 1$ формула не верна.

3.4.2*. Одинаковость испытаний

Модель Бернулли описывает опыты, составленные из *одинаковых* испытаний. Одинаковость испытаний выражается в том, что при каждом испытании шансы на успех равны и возможности неудачи одинаковы. Классическим примером одинаковых испытаний являются подбрасывания симметричной монеты. В соответствии с предположением об одинаковости испытаний в модели Бернулли при каждом испытании вероятность успеха равна a, а вероятность неудачи $1-a$. Это выражает вытекающее из формулы успехов и неудач

Следствие 1. $P(У_i) = a$, $P(H_i) = (1-a)$ $(i = 1, \ldots, n)$.

Доказательство. Эти равенства получаются из формулы успехов и неудач соответственно при $K = \{i\}$, $L = 0$ $(k=1,\ l=0)$ и $K = 0$, $L = \{i\}$ $(k=0,\ l=1)$.

Таким образом, в модели Бернулли одинаковость испытаний выражается в том, что в каждом испытании

$$\mathscr{B}_i = \{O, У_i, H_i, U\}$$

вероятности соответствующих событий равны

$$P(O)=0, \quad P(У_i)=a, \quad P(H_i)=1-a, \quad P(U)=1$$

для каждого номера $i=1, \ldots, n$.

4.4.2*. Независимость испытаний

Модель Бернулли описывает опыты, составленные из *независимых* испытаний. Независимость испытаний выражается в том, что шансы на получение данных результатов при данных испытаниях не зависят от результатов любых других испытаний. Классическим примером независимых испытаний являются подбрасывания монеты. В соответствии с предположением о независимости испытаний в модели Бернулли вероятность пересечения произвольных исходов каждых испытаний равна произведению вероятностей этих исходов. Это выражает вытекающее из формулы успехов и неудач

Следствие 2. $P\left(\bigcap\limits_{i\in K} У_i \bigcap\limits_{j\in L} H_j\right)=\prod\limits_{i\in K} P(У_i) \prod\limits_{j\in L} P(H_j)$.

Доказательство. Обе части равны числу $a^k(1-a^l)$.

Пример. $P(У_2У_4H_6)=P(У_2)P(У_4)P(H_6)=a^2(1-a)$.

Замечание. В следствии 2 также существенно условие $KL=O$.

5.2. Вероятностная модель Бернулли

Множество исходов $U=B^n$, алгебра событий $\mathscr{A}=\mathscr{B}^n$, элементарная вероятность $p=q^n$ и вероятность $P=Q^n$ образуют *модель Бернулли n испытаний с вероятностью успеха a.* Эта модель описывает опыт, составленный из *последовательности n одинаковых и независимых испытаний,* каждое из которых оканчивается *успехом* либо *неудачей.* Причем в каждом испытании шансы на успех оцениваются числом a, а возможность неудачи — числом $1-a$.

§ 3*. СВОЙСТВА ВЕРОЯТНОСТИ

Общие свойства элементарной вероятности и вероятности в модели Бернулли те же, что и в модели Лапласа: элементарная вероятность p является *положительной нормированной функцией на множестве исходов*, а вероятность P — *положительной нормированной аддитивной функцией на классе событий.*

Рассмотрим модель Бернулли n испытаний с вероятностью успеха a.

1.3. Элементарная вероятность p

1. Элементарная вероятность p *положительна.* Это следует из неравенств $0\leqslant a\leqslant 1$ и равенства (3) § 2:

$$p(u)=a^{s(u)}(1-a)^{n-s(u)}\geqslant 0$$

для каждого исхода u. Символ $s(u)$ обозначает сумму элементов строки u:

$$s(u) = \sum_{1 \leqslant j \leqslant n} u_j.$$

Эта сумма равна числу единиц в строке u.

В частности, для $n=2$ элементарную вероятность определяет следующая

Таблица элементарной вероятности

u	$s(u)$	$p(u)$
11	2	$a^2 = a^2(1-a)^{2-2}$
10	1	$a(1-a) = a^1(1-a)^{2-1}$
01	1	$a(1-a) = a^1(1-a)^{2-1}$
00	0	$(1-a)^2 = a^0(1-a)^{2-0}$

2. Элементарная вероятность p *нормирована*. Это следует из правила прямоугольника:

$$\sum_{u \in U} p(u) = P(U) = [Q(B)]^n = 1^n = 1.$$

2.3. Вероятность P

1. Вероятность P *положительна*:

$$P(A) = \sum_{u \in A} p(u) \geqslant 0$$

для каждого события A. Это следует из положительности элементарной вероятности p.

2. Вероятность P *нормирована*:

$$P(U) = \sum_{u \in U} p(u) = 1.$$

Это эквивалентно нормированности элементарной вероятности p.

3. Вероятность P *аддитивна*:

$$P(A+B) = P(A) + P(B)$$

для каждых непересекающихся событий A и B. Это правило сложения для вероятности P следует из ассоциативности сложения для чисел:

$$P(A+B) = \sum_{u \in A+B} p(u) = \sum_{u \in A} p(u) + \sum_{u \in B} p(u) = P(A) + P(B) \quad (AB = 0).$$

Пример. Если $n=2$, $A=\{11, 10\}$, $B=\{01, 00\}$, то

$$AB = 0, \quad A+B = U$$

и

$$P(A+B)=p(11)+p(10)+p(01)+p(00),$$
$$P(A)=p(11)+p(10),\quad P(B)=p(01)+p(00),$$
$$P(A+B)=P(A)+P(B).$$

3.3. Примеры

Рассмотрим несколько простых примеров вычисления вероятностей событий в модели Бернулли. Эти вычисления повторяют, по существу, некоторые вычисления примеров 3 и 4 из § 1.

Пример 1. *Симметричная монета подбрасывается два раза. Какова вероятность того, что:*

1) при первом подбрасывании появляется герб?

2) при первом подбрасывании появляется цифра?

3) при втором подбрасывании появляется герб?

4) при втором подбрасывании появляется цифра?

Решим эту задачу, используя модель Бернулли $n=2$ испытаний с вероятностью успеха $a=1/2$, рассматривавшуюся в примере 3 § 1. Она сводится к вычислению вероятностей событий $У_1$, H_1, $У_2$, H_2. По формуле успехов и неудач находим:

$$P(У_1)=P(У_2)=a=1/2,\quad P(H_1)=P(H_2)=1-a=1/2.$$

Пример 2. *Производится два выстрела по цели. Какова вероятность того, что:*

1) при первом выстреле происходит попадание?

2) при первом выстреле происходит промах?

3) при втором выстреле происходит попадание?

4) при втором выстреле происходит промах?

Известно, что при большом числе выстрелов, произведенных в таких же условиях, частота попадания была равна 9/10.

Решим эту задачу, используя модель Бернулли $n=2$ испытаний с вероятностью успеха $a=9/10$, рассматривавшуюся в примере 4 § 1. Она сводится к вычислению тех же событий $У_1$, H_1, $У_2$, H_2, что и в примере 1. Как и в этом примере, по формуле успехов и неудач получаем:

$$P(У_1)=P(У_2)=a=9/10,\quad P(H_1)=P(H_2)=1-a=1/10.$$

Пример 3. *Симметричная монета подбрасывается два раза. Какова вероятность того, что:*

1) герб не появится ни разу?

2) герб появится ровно один раз?

3) герб появится два раза?

Решим эту задачу, используя модель Бернулли $n=2$ испытаний с вероятностью успеха $a=1/2$, рассматривавшуюся в примере 3 § 1. Она сводится к вычислению вероятностей событий:

$$A_0=\{u:s(u)=0\}=\{00\}=H_1H_2,$$
$$A_1=\{u:s(u)=1\}=\{10,01\}=У_1H_2+H_1У_2,$$
$$A_2=\{u:s(u)=2\}=\{11\}=У_1У_2.$$

Используя правило сложения вероятностей и формулу успехов и неудач, получаем:

$$P(A_0)=(1-a)^2=1/4, \quad P(A_1)=P(У_1Н_2)+P(Н_1У_2)=$$
$$=2a(1-a)=1/2, \quad P(A_2)=a^2=1/4.$$

Заметим, что события A_0, A_1, A_2 попарно не пересекаются, и поэтому

$$P(U)=P(A_0+A_1+A_2)=P(A_0)+P(A_1)+P(A_2)=$$
$$=(1-a)^2+2a(1-a)+a^2=((1-a)+a)^2=1.$$

Пример 4. *Производится два выстрела по цели. Какова вероятность того, что:*

1) попадания не происходит ни разу?

2) попадание происходит ровно один раз?

3) попадание происходит два раза?

Известно, что при большом числе выстрелов, произведенных в таких же условиях, частота попадания была равна 9/10.

Решим эту задачу, используя модель Бернулли $n=2$ испытаний с вероятностью исхода $a=9/10$, рассматривавшуюся в примере 4 § 1. Она сводится к вычислению вероятностей тех же событий A_0, A_1, A_2, что и в примере 3. Те же вычисления, что и в этом примере, дают:

$$P(A_0)=(1-a)^2=1/100, \quad P(A_1)=2a(1-a)=18/100,$$
$$P(A_2)=a^2=81/100.$$

Точно так же, как и в примере 3, верны равенства

$$P(U)=P(A_0+A_1+A_2)=P(A_0)+P(A_1)+P(A_2)=1.$$

§ 4*. УСЛОВНАЯ ВЕРОЯТНОСТЬ

Содержание понятия условной вероятности в модели Бернулли то же, что и в модели Лапласа. Пример 3 с подбрасыванием симметричной монеты из § 4 главы 1 является примером и для модели Бернулли.

1.4. Определения

Рассмотрим модель Бернулли n испытаний с вероятностью успеха a. Условимся называть событие B *возможным*, если его вероятность не равна нулю:

$$P(B)>0.$$

Заметим, что если $a=0$ или $a=1$, то некоторые события, не равные 0, не являются возможными. Если же $0<a<1$, то единственным не возможным событием является пустое событие O, как и в модели Лапласа.

Рассмотрим произвольное возможное событие B. Определим *условную элементарную вероятность* p_B как функцию на множестве исходов U, значение $p(u)$ которой для каждого исхода u определяется равенствами:

(1) $$p_B(u) = p(u)/P(B) \quad (u \in B),$$

(2) $$p_B(u) = 0 \quad (u \notin B).$$

В этом определении используется предполагаемая возможность события B, обеспечивающая неравенство $P(B) \neq 0$.

Заметим, что из равенства (3) § 2 для p следует, что равенство (1) эквивалентно равенству

(3) $$p_B(u) = a^{s(u)}(1-a)^{n-s(u)}/P(B),$$

где $s(u)$ обозначает сумму элементов строки u:

$$s(u) = \sum_{1 \leqslant j \leqslant n} u_j$$

и равно числу единиц в этой строке.

В примере 3 из § 4 главы 1 приведены значения условной вероятности p_B для случая $n=2$, $a=1/2$ и $B=\{10, 11\}$.

Упражнение. Вычислить значения условной вероятности P_B для случая $n=3$, $0<a<1$ и $B=\{111, 110, 101, 100\}$.

Определим *условную вероятность* P_B как функцию на классе событий $\mathscr{A}$, значение $P_B(A)$ которой для каждого события A определяется равенством

(4) $$P_B(A) = \sum_{u \in A} p_B(u).$$

В модели Бернулли, так же как и в модели Лапласа, условная вероятность P_B является мерой реализуемости явлений, определяемых результатами рассматриваемого опыта, *при условии*, что реализуется явление, описываемое событием B.

2.4. Правила умножения и деления

Рассмотрим модель Бернулли n испытаний с вероятностью успеха a. Для каждых события A и возможного события B в этой модели верно равенство

(5) $$P_B(A) = P(AB)/P(B).$$

В самом деле, используя равенства (1), (2) и (4) для условных элементарной вероятности и вероятности, получаем:

$$P_B(A) = \sum_{u \in A} p_B(u) = \sum_{u \in AB} p_B(u) = \sum_{u \in AB} p(u)/P(B) = P(AB)/P(B).$$

Из равенства (5) следует равенство

(6) $$P(AB) = P(B)P_B(A).$$

Равенство (6) называется *правилом умножения*, а равенство (5) — *правилом деления* для вероятности.

Так же, как и в модели Лапласа, в модели Бернулли правила умножения и деления используются для вычисления вероятностей событий.

3.4. Примеры

Рассмотрим модель Бернулли $n=2$ испытаний с вероятностью успеха $a=1/2$, описанную в примере 3 § 1.

Пример 1. Из равенств (2″) примера 3 и правила деления вытекают равенства:

$$P_{У_1}(У_2) = P(У_2), \quad P_{У_1}(H_2) = P(H_2),$$
$$P_{H_1}(У_2) = P(У_2), \quad P_{H_1}(H_2) = P(H_2).$$

Эти равенства соответствуют тому факту, что результат второго подбрасывания монеты не зависит от результата первого подбрасывания.

Пример 2. Из равенств (2″) примера 3 и правила деления вытекают равенства:

$$P_{У_2}(У_1) = P(У_1), \quad P_{У_2}(H_1) = P(H_1),$$
$$P_{H_2}(У_1) = P(У_1); \quad P_{H_2}(H_1) = P(H_1).$$

Эти равенства соответствуют тому факту, что результат первого подбрасывания монеты не зависит от результата второго подбрасывания.

Замечание. Примеры 1 и 2 позволяют сказать, что равенства

$$P(У_1У_2)=P(У_1)P(У_2), \quad P(У_1H_2)=P(У_1)P(H_2),$$
$$P(H_1У_2)=P(H_1)P(У_2), \quad P(H_1H_2)=P(H_1)P(H_2)$$

примера 3 из § 1 выражают независимость подбрасываний монеты.

Пример 3. *Симметричная монета подбрасывается два раза. Известно, что хотя бы один раз появляется герб. Какова вероятность при этом условии того, что два раза появляется герб?*

В рассматриваемой модели задача сводится к вычислению условной вероятности $P_B(A)$ события $A=\{11\}$ при условии $B=\{11, 10, 01\}$. Используя правило деления, получаем:

$$P_B(A)=P(AB)/P(B)=1/3.$$

Если вернуться к игре с монетой, описанной в примере 3 § 1 главы 1, то полученная вероятность $P_B(A)$ является условной вероятностью выигрыша при условии, что игрок не проигрывает. Эта вероятность больше вероятности выигрыша:

$$P(A)=1/2\cdot 1/2=1/4.$$

§ 5*. НЕЗАВИСИМОСТЬ И ЗАВИСИМОСТЬ

Содержание понятий *независимости* и *зависимости* в модели Бернулли то же, что и в модели Лапласа. Основную роль в модели

Бернулли играет *независимость испытаний*. Данный параграф посвящен уточнению этого понятия.

Рассмотрим модель Бернулли n испытаний с вероятностью успеха a.

1.5. Независимость двух событий

Назовем события A и B *независимыми*, если верно равенство

$$P(AB)=P(A)P(B). \tag{7}$$

Если равенство (7) не верно, то будем называть события A и B *зависимыми*.

В примерах 1 и 2 из § 4 события $У_1$ и $У_2$, H_1 и $У_2$, $У_1$ и H_2, H_1 и H_2 независимы. В примере 3 из § 4 события A и B зависимы:

$$P(AB)=1/4\neq 1/4\cdot 3/4=P(A)P(B).$$

Если события A и B *возможны*:

$$P(A)>0, \quad P(B)>0,$$

то равенство (7) эквивалентно каждому из равенств

$$P_B(A)=P(A), \quad P_A(B)=P(B).$$

Это следует из равенств (5) и (6).

В примере 3 из § 4:

$$P_B(A)=1/3\neq 1/4=P(A),$$
$$P_A(B)=1\neq 3/4=P(B).$$

2.5. Независимость семейства событий

Назовем события C_i семейства $(C_i)_{i\in I}$ с конечным множеством индексов *независимыми*, если для каждого множества $M\subseteq I$ верно равенство

$$P\left(\bigcap_{i\in M} C_i\right)=\prod_{i\in M} P(C_i). \tag{8}$$

Если для некоторого множества $K\subseteq I$ равенство (8) не верно, то будем называть события семейства $(C_i)_{i\in I}$ *зависимыми*.

Пример 1. События $C_i=У_i$ семейства $(C_i)_{i\in\{1,\ldots,n\}}$ независимы, так как для каждого множества $M\subseteq\{1,\ldots,n\}$ верно равенство

$$P\left(\bigcap_{i\in M} У_i\right)=\prod_{i\in M} P(У_i).$$

Это вытекает из следствия 2 формулы успехов и неудач.

Пример 2. События $C_1=У_1$, $C_2=H_1$, $C_3=H_3$ семейства $(C_i)_{i\in\{1,2,3\}}$ зависимы, если $0<a<1$, так как в этом случае для множества $M=\{1, 2\}$ равенство (8) не верно:

$$P(У_1H_1)=P(O)=0\neq a(1-a)=P(У_1)P(H_1).$$

Определение независимости семейства событий позволяет придать точный смысл предложению: *исходы различных испытаний в модели Бернулли независимы.*

Рассмотрим произвольное семейство $(B_i)_{i=1,\ldots,n}$ исходов B_i испытаний $i=1,\ldots,n$. Обозначим множество номеров i, для которых $B_i=У_i$, буквой S, а множество номеров j, для которых $B_j=H_j$,— буквой T:

$$ST=O, \quad S+T=\{1,\ldots,n\}.$$

Пример. $\{1,\ldots,n\}=\{1, 2, 3, 4, 5, 6\}$,

$$S=\{2, 4, 6\}, \quad T=\{1, 3, 5\},$$

$$B_1=H_1, \quad B_2=У_2, \quad B_3=H_3,$$

$$B_4=У_4, \quad B_5=H_5, \quad B_6=У_6.$$

Лемма. *Исходы B_i $(i=1,\ldots,n)$ независимы.*

Доказательство. Рассмотрим произвольное конечное множество $M\subseteq\{1,\ldots,n\}$ и множества $K=MS$, $L=MT$. Заметим, что

$$KL=O, \quad K+L=M.$$

По следствию 2 формулы успехов и неудач получаем:

$$P\left(\bigcap_{i\in M} B_i\right)=P\left(\bigcap_{i\in K} У_i \bigcap_{j\in L} H_j\right)=\prod_{i\in K} P(У_i)\prod_{j\in L} P(H_j)=$$

$$=\prod_{i\in K} P(B_i)\prod_{j\in L} P(B_j)=\prod_{j\in M} P(B_i).$$

Значит, события B_i семейства $(B_i)_{i=1,\ldots,n}$ независимы.

Лемма доказана.

Пример. $S=\{2, 4, 6\}$, $T=\{1, 3, 5\}$,

$$M=\{1, 2, 3\}, \quad K=\{1, 2, 3\}\cdot\{2, 4, 6\}=\{2\},$$

$$L=\{1, 2, 3\}\cdot\{1, 3, 5\}=\{1, 3\},$$

$$P(B_1B_2B_3)=P(У_2H_1H_3)=P(У_2)P(H_1)P(H_3)=$$

$$=P(B_2)P(B_1)P(B_3).$$

3.5. Независимость семейства классов

Назовем классы $\mathscr{C}_i$ событий семейства $(\mathscr{C}_i)_{i\in I}$ с конечным множеством индексов *независимыми*, если события $C_i\in\mathscr{C}_i$ каждого семейства $(C_i)_{i\in I}$ независимы. Если события $C_i\in\mathscr{C}_i$ некоторого семейства $(C_i)_{i\in I}$ зависимы, то классы $\mathscr{C}$ событий будем называть *зависимыми.* Другими словами, классы $\mathscr{C}_i$ независимы, если как бы ни выбирать по одному событию C_i из каждого класса $\mathscr{C}_i$, эти события оказываются независимыми. Если же можно таким образом выбрать зависимые события C_i, то классы $\mathscr{C}_i$ считаются зависимыми.

Пример 1. Испытания $\mathscr{B}_1=\{O, У_1, H_1, U\}$, $\mathscr{B}_2=\{O, У_2, H_2, U\}$ семейства $(\mathscr{B}_i)_{i=1,2}$ независимы, так как равенство

$$P(B_1B_2)=P(B_1)P(B_2)$$

верно для каждых событий $B_1\in\mathscr{B}_1$ и $B_2=\mathscr{B}_2$.

Пример 2. Если $0<a<1$, то классы

$$\mathscr{C}_1=\{O,\ У_1,\ H_1,\ U\},\quad \mathscr{C}_2=\{У_1,\ У_2\}$$

семейства $(\mathscr{C}_i)_{i=1,2}$ зависимы, так как события $C_1=У_1\in\mathscr{C}_1$ и $C_2=$ $=У_1\in\mathscr{C}_2$ зависимы:

$$P(У_1У_1)=P(У_1)=a\neq a^2=P(У_1)P(У_1).$$

Заметим, что любые другие два события $C_1\in\mathscr{C}_1$ и $C_2\in\mathscr{C}_2$ независимы.

Определение независимых классов событий позволяет придать точный смысл предложению: *различные испытания в модели Бернулли независимы.*

Теорема. *Испытания* $\mathscr{B}_i$ $(i=1,\ldots,n)$ *независимы.*

Доказательство. Из леммы о независимости исходов следует, что классы

$$\mathscr{C}_i=\{У_i,\ H_i\}$$

семейства $(\mathscr{C}_i)_{i=1,\ldots,n}$ независимы.

Заметим, что если $C_i=O$ для некоторого номера $i\in M$, то в равенстве (8) обе части равны нулю. А если $C_j=U$ для некоторого номера $j\in M$, то равенство (8) сводится к аналогичному равенству для множества номеров $M-\{j\}$. Поэтому из независимости классов $\mathscr{C}_i$ семейства $(\mathscr{C}_i)_{i=1,\ldots,n}$ следует независимость классов

$$\mathscr{B}_i=\{O,\ У_i,\ H_i,\ U\}$$

семейства $(\mathscr{B}_i)_{i=1,\ldots,n}$.

Теорема доказана.

Задачи, поясняющие использование модели Бернулли, собраны в главе 1 части III.

Глава 3

КОНЕЧНЫЕ ВЕРОЯТНОСТНЫЕ МОДЕЛИ

Игрок следит за подброшенной монетой. Она падает гербом вверх: успех, ставка выиграна. Монета подбрасывается еще раз и падает цифрой вверх: неудача, ставка проиграна. Монета подбрасывается снова и снова. Успехи чередуются с неудачами, причем результат каждого подбрасывания определяется случаем. Но, если игра продолжается достаточно долго, то, как правило, число выигрышей оказывается примерно равным числу проигрышей. В этой закономерности проявляется симметричность монеты.

Стрелок прицеливается, стреляет и... не попадает в цель. Неудача. Он прицеливается тщательнее, стреляет еще раз: успех, цель поражена. Стрелок производит выстрел за выстрелом. Попадания чередуются с промахами, причем результат каждого вы-

стрела в определенной степени зависит от случая. Но, если стрельба продолжается достаточно долго, то, как правило, частота попаданий оказывается близкой к наблюдавшейся раньше. В этом устойчивом поведении частоты попаданий проявляется квалификация стрелка.

Контролер проверяет партию одинаковых изделий и выбирает наугад одно из них. Оно оказывается годным. С точки зрения контролера, целью которого является обнаружение брака, это неудача. Второе, третье, ..., десятое наугад выбранные изделия тоже годные. Наконец, одиннадцатое изделие оказывается негодным: успех. Контроль продолжается, годные изделия чередуются с негодными, причем результат каждой проверки зависит от случая. Но, если число выбираемых для проверки изделий достаточно велико, то, как правило, доля негодных среди них оказывается близкой к доле негодных изделий во всей проверяемой партии.

Многие задачи, связанные с играми, стрельбой по цели, контролем качества и другими процессами, решаются с помощью простых вероятностных моделей.

Каждая вероятностная модель составляется из *множества исходов*, *алгебры событий и вероятности.* В этой главе описываются модели, множества исходов которых конечны.

Множество исходов описывает возможные результаты проводимого опыта или наблюдения: падение гербом или цифрой вверх подбрасываемой монеты, попадание или промах стрелка, годность или негодность проверяемого изделия. Формально множеством исходов может быть каждое непустое конечное множество U.

Алгебра событий описывает явления, связанные с проводимым опытом или наблюдением и определяемые его результатами: двукратное падение гербом вверх дважды подбрасываемой монеты, хотя бы одно попадание при двух выстрелах, годность всех проверяемых изделий. Алгебраические действия с событиями соответствуют логическим связкам: двукратное появление герба означает появление герба при первом *и* при втором подбрасывании монеты, хотя бы одно попадание при двух выстрелах означает попадание при первом *или* при втором выстреле; годность всех проверяемых изделий означает, что *не* оказывается хотя бы одного негодного изделия. Формально алгеброй событий может быть каждая алгебра $\mathcal{A}$ частей конечного множества исходов U.

Вероятность служит мерой реализуемости рассматриваемых явлений. Вероятность каждого события определяется как сумма *элементарных вероятностей* составляющих это событие исходов. Элементарная вероятность каждого исхода измеряет реализуемость описываемого этим исходом возможного результата и выбирается в соответствии с условиями задачи. В частности, если все результаты равновозможны, то им приписывается одинаковая элементарная вероятность, равная отношению единицы к числу всех возможных результатов. Вероятность каждого события в этой классической модели равна отношению числа составляющих событие исходов к числу всех возможных исходов. Формально эле-

ментарной вероятностью может быть каждая положительная функция p на множестве исходов U, сумма значений которой равна единице. Вероятность P является функцией на классе событий $\mathscr{A}$. Значения вероятности P являются суммами значений элементарной вероятности p.

Каждая задача в такой модели формулируется как задача о вычислении вероятности некоторого события. Для решения этой задачи используются различные правила вычисления вероятностей: правило сложения, правило умножения, формула полной вероятности. Рассматриваемая простая модель позволяет решить много интересных задач с самым различным содержанием.

§ 1. ПРИМЕРЫ

1.1. Пример 1. Выстрел по мишени

Рассмотрим опыт, заключающийся в выстреле по стандартной мишени, и построим вероятностную модель этого опыта.

1.1.1. Множество исходов U

Возможными результатами рассматриваемого опыта являются промах (0 очков) или получение 1, 2, ..., 10 очков. Условимся эти результаты соответственно описывать числами 0, 1, 2, ..., 10 и называть эти числа исходами. Составленное ими множество исходов $U=\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ описывает возможные результаты выстрела по мишени.

2.1.1. Алгебра событий $\mathscr{A}$

Каждую часть A множества исходов U условимся называть событием. Алгебру $(\mathscr{A}, \cup, \cap, ')$ образованную классом $\mathscr{A}$ всех событий и их объединением, пересечением и дополнением, назовем алгеброй событий. Сокращенно будем обозначать ее буквой $\mathscr{A}$ (так же, как и класс событий). Алгебра событий $\mathscr{A}$ описывает явления, определяемые возможными результатами рассматриваемого опыта, и логические связи между этими явлениями.

Например, событие $A=\{6, 7, 8, 9, 10\}$ описывает получение числа очков строго большего пяти. Элементарное событие $B=\{0\}$ описывает промах, его дополнение $C=\{1, 2, \ldots, 10\}$ — попадание.

3.1.1. Элементарная вероятность p

Предположим, что в тех же условиях прежде было произведено большое число выстрелов и частота $\nu(u)$ получения u очков выражается в процентах следующей таблицей:

u	0	1	2	3	4	5	6	7	8	9	10
$\nu(u)$	1	1	1	1	1	5	5	5	10	20	50

Одинаковость условий позволяет использовать частоту $\nu(u)$ для оценки шансов на получение u очков при описываемом выстреле.

В соответствии с этим рассмотрим функцию p на множестве исходов U со значением $p(u)=\nu(u)$ для каждого исхода u. Назовем функцию p элементарной вероятностью. Значение $p(u)$ описывает шансы на получение u очков при выстреле.

Заметим, что элементарная вероятность p, равная частоте ν, наблюдавшейся прежде, является вместе с этой частотой положительной нормированной функцией на множестве исходов U:

$$p(u)\geqslant 0, \quad \dot{\Sigma} p(u)=1.$$

4.1.1. Вероятность P

Для каждого события A вероятность $P(A)$ определим как сумму элементарных вероятностей $p(u)$ исходов u, составляющих событие A:

$$P(A)=\sum_{u\in A} p(u).$$

Вероятность $P(A)$ служит мерой реализуемости явления, описываемого событием A. Функцию P на классе событий $\mathscr{A}$ со значением $P(A)$ для каждого события A назовем вероятностью.

Множество исходов U, алгебра событий $\mathscr{A}$ и вероятность P образуют удобную модель выстрела по мишени.

С помощью построенной вероятностной модели решается, например, такая

Задача. *Какова вероятность того, что при выстреле по мишени получается число очков строго большее пяти?*

Решение. Задача сводится к вычислению вероятности $P(A)$ события $A=\{6, 7, 8, 9, 10\}$. Используя таблицу, находим, что

$$P(A)=p(6)+p(7)+p(8)+p(9)+p(10)=$$
$$=(5+5+10+20+50)/100=9/10.$$

Легко вычислить также вероятности промаха и попадания:

$$P(B)=p(0)=1/100, \quad P(C)=1-P(B)=99/100.$$

Замечание. Использовать модель Лапласа для описания выстрела по мишени, по-видимому, невозможно: совершенно не понятно, как выбрать равновозможные результаты. Если интересоваться только попаданием и промахом, а не числом очков, то можно использовать модель Бернулли $n=1$ испытания с вероятностью успеха $a=99/100$.

2.1. Пример 2. Выбор буквы из книги

Рассмотрим опыт, заключающийся в выборе наугад буквы из книги, написанной на русском языке, и построим вероятностную модель этого опыта.

1.2.1. Множество исходов U

В качестве множества исходов U удобно взять алфавит из 32 букв:

$\{U = \{$а, б, в, г, д, е, ж, з, и, й, к, л, м, н, о, п, р, с, т, у, ф, х, ч, ш, щ, ь, ы, э, ю, я, —$\}$.

В этом алфавите буква «е» обозначает также букву «ё», буква «ь» — букву «ъ». Черточка «—» — это пробел между словами. Условимся называть пробел также *пустой* буквой.

2.2.1. Алгебра событий $\mathcal{A}$

В качестве алгебры событий удобно взять алгебру $\mathcal{A}$, образованную классом $\mathcal{A}$ всех частей множества исходов U и объединением, пересечением и дополнением для $\mathcal{A}$.

Примером события является множество $A = \{$а, е, и, о, у, ы, э, ю, я$\}$, описывающее выбор гласной буквы.

3.2.1. Элементарная вероятность p

Если рассматривается книга не слишком специальная, то для оценки шансов на выбор буквы «и» естественно взять частоту $\nu(u)$ буквы «и» в книгах, написанных на русском языке. Такие частоты собраны в следующей таблице (А. М. Яглом, И. М. Яглом. «Вероятность и информация». М., «Наука», 1973):

— 0,175	о 0,090	е, ё 0,072	а 0,062	и 0,062	т 0,053	н 0,053	с 0,045
р 0,040	в 0,038	л 0,035	к 0,028	м 0,026	д 0,025	п 0,023	у 0,021
я 0,018	ы 0,016	з 0,016	ь, ъ 0,014	б 0,014	г 0,013	ч 0,012	й 0,010
х 0,009	ж 0,007	ю 0,006	ш 0,006	ц 0,004	щ 0,003	э 0,003	ф 0,002

В соответствии с этим в качестве элементарной вероятности удобно взять функцию p на множестве исходов U со значением $p(u) = \nu(u)$ для каждого исхода u.

Заметим, что элементарная вероятность p, равная частоте ν, является вместе с этой частотой положительной нормированной функцией на множестве исходов U:

$$p(u) \geqslant 0, \quad \Sigma p(u) = 1.$$

4.2.1. Вероятность P

Вероятность P определяется как функция P на классе событий $\mathscr{A}$ со значением

$$P(A) = \sum_{u \in A} p(u)$$

для каждого события A. Вероятность служит мерой реализуемости явления, описываемого событием A.

Множество исходов U, алгебра событий $\mathscr{A}$ и вероятность P образуют удобную модель выбора наугад буквы из книги.

С помощью построенной вероятностной модели решается, например, такая

Задача. *Какова вероятность того, что выбранная наугад из книги буква является гласной?*

Решение. Задача сводится к вычислению вероятности $P(A)$ события $A = \{$а, е, и, о, у, ы, э, ю, я$\}$. Используя таблицу частот, находим, что

$$P(A) = p(\text{а}) + p(\text{е}) + p(\text{и}) + p(\text{о}) + p(\text{у}) + \\ + p(\text{ы}) + p(\text{э}) + p(\text{ю}) + p(\text{я}) = 0{,}062 + 0{,}072 + \\ + 0{,}062 + 0{,}090 + 0{,}021 + 0{,}016 + 0{,}003 + 0{,}006 + \\ + 0{,}018 = 0{,}350.$$

Замечание. Использовать модель Лапласа для описания выбора буквы из книги, по-видимому, невозможно. Если интересоваться только гласностью буквы, то можно использовать модель Бернулли $n = 1$ испытания с вероятностью успеха $a = 35/100$.

Упражнение. Составить аналогичную таблицу частот для 1000 букв, взятых подряд из какой-нибудь книги, и сравнить с приведенной в 3.2.1.

3.1. Пример 3. Простые числа

Разобьем множество натуральных чисел $\{1, \ldots, 100\}$ на десятки $\{1, \ldots, 10\}$, $\{11, \ldots, 20\}, \ldots, \{91, \ldots, 100\}$ и предположим, что нас интересует

Задача. *Какова вероятность того, что в наугад выбранном десятке оказывается нечетное количество простых чисел?*

Решение 1. Выбор наугад позволяет использовать для решения задачи стандартную модель Лапласа 10 равновероятных исходов $k = 1, \ldots, 10$. Исход u описывает выбор u-го в указанном порядке

десятка. Для того, чтобы определить событие, описывающее интересующее нас появление нечетного количества простых чисел, выпишем все простые числа из множества $\{1, \ldots, 100\}$, разбив их по десяткам, которым они принадлежат:

1	2	3	4	5	6	7	8	9	10
2, 3, 5, 7	11, 13, 17, 19	23, 29	31, 37	41, 43, 47	53, 59	61, 67	71, 73, 79	83, 89	97

Из таблицы видно, что задача сводится к вычислению вероятности $P(A)$ события $A = \{5, 8, 10\}$:

$$P(A) = n(A)/n(U) = 3/10.$$

Решение 2. Таблица из решения 1 позволяет использовать для решения задачи и другую вероятностную модель.

1.3.1. Множество исходов U

Возьмем в качестве множества исходов множество $U = \{1, 2, 3, 4\}$. Исход u описывает выбор десятка, в котором ровно u простых чисел.

2.1.3. Алгебра событий $\mathscr{A}$

Как обычно, возьмем в качестве алгебры событий алгебру, образованную классом $\mathscr{A}$ всех частей множества исходов U и объединением, пересечением и дополнением этих частей. Интересующее нас появление нечетного количества простых чисел описывает событие $A = \{1, 3\}$.

3.1.3. Элементарная вероятность p

Для оценки шансов на появление u простых чисел в выбираемом десятке естественно использовать долю $\nu(u)$ десятков, которым принадлежит ровно u простых чисел. В соответствии с этим значения элементарной вероятности p определим равенствами

$$p(1) = 1/10, \quad p(2) = 5/10, \quad p(3) = 2/10, \quad p(4) = 2/10.$$

4.1.3. Вероятность P

Значение $P(A)$ вероятности P для каждого события A определим равенством

$$P(A) = \sum_{u \in A} p(u).$$

В модели, образованной так определенными множеством исходов U, алгеброй событий $\mathscr{A}$ и вероятностью P, задача сводится к

вычислению вероятности $P(A)$ события $A=\{1, 3\}$:

$$P(A)=\sum_{u \in A} p(u)=p(1)+p(3)=1/10+2/10=3/10.$$

Таким образом, оба решения приводят к одному и тому же ответу.

§ 2. КОНЕЧНАЯ ВЕРОЯТНОСТНАЯ МОДЕЛЬ

Для описания опытов с конечным множеством возможных результатов удобна вероятностная модель, определяемая произвольным конечным множеством U и положительной нормированной функцией p на нем.

1.2. Множество исходов U

Основой модели является множество исходов U.

Определение 1. *Множеством исходов называется каждое непустое конечное множество* U.

Элементы u множества исходов U называются исходами.

Пример 1. Модель Лапласа: $U=\{1, \ldots, n\}$.

Пример 2. Модель Бернулли: $U=B^n$.

Пример 3. В примерах 1—3 из § 1 рассматривались множества исходов $U=\{1, \ldots, 10\}$, {а, ..., я}, $\{1, \ldots, 4\}$.

Содержание. Множество исходов U описывает возможные *результаты опыта*. Примеры 1—3 из § 1 поясняют это описание.

2.2. Алгебра событий $\mathscr{A}$

Вторым элементом модели является алгебра событий $\mathscr{A}$.

Определение 2. *Событием называется каждая часть A множества исходов* U.

Класс $\mathscr{A}$ всех событий, их объединение $\cup$, пересечение $\cap$ и дополнение $'$ образуют алгебру событий $(\mathscr{A}, \cup, \cap, ')$. Сокращенно алгебра событий обозначается буквой $\mathscr{A}$ (так же, как и класс).

Пример 1. Пустое множество O называется невозможным событием.

Пример 2. Множество исходов U называется достоверным событием.

Пример 3. В примерах 1—3 из § 1 рассматривались события $A=\{6, 7, 8, 9, 10\}$, {а, е, и, о, у, ы, э, ю, я}, $\{1, 3\}$.

Содержание. Алгебра событий $\mathscr{A}$ описывает ***явления***, определяющиеся возможными результатами опыта. Примеры 1—3 из § 1 поясняют это описание.

Замечание. Предлагаемое стандартное определение алгебры событий достаточно для всех приложений, рассматриваемых в этой книге. Можно было бы определить события как некоторые части множества исходов U, класс которых с объединением, пересечением и дополнением этих частей образует алгебру.

Рассмотрим, например, опыт, состоящий из последовательности двух испытаний, каждое из которых имеет два возможных ре-

зультата. Предположим, что нас интересуют результаты только второго испытания и логические связи между ними. Возможные результаты опыта описывает множество исходов $U=\{11, 10, 01, 00\}$, результаты второго испытания — множества $У_2=\{11, 01\}$ и $H_2=\{10, 00\}$. Класс $\mathscr{B}_2=\{O, У_2, H_2, U\}$ с объединением, пересечением и дополнением множеств его образует алгебру. В рассматриваемом случае событиями можно было бы назвать множества O, $У_2$, H_2, U. Остальные части множества U (в частности, $\{11\}$, $\{10\}$, $\{01\}$, $\{00\}$) не считались бы событиями.

3.2. Элементарная вероятность p

Третьим элементом модели является элементарная вероятность p.

Определение 3. *Элементарной вероятностью называется каждая положительная нормированная функция p на множестве исходов U.*

Значение $p(u)$ элементарной вероятности p для каждого исхода u называется элементарной вероятностью исхода u.

Положительность элементарной вероятности p означает, что элементарная вероятность $p(u)$ каждого исхода u положительна:

$$p(u) \geqslant 0.$$

Нормированность элементарной вероятности p означает, что сумма элементарных вероятностей $p(u)$ для всех исходов u равна единице:

$$\sum_{u \in U} p(u) = 1.$$

Пример 1. Модель Лапласа: $p(u)=1/n(U)$.

Пример 2. Модель Бернулли: $p(u)=a^{s(u)}(1-a)^{n-s(u)}$.

Пример 3. В примерах 1—3 из § 1 рассматривались элементарные вероятности, определяемые приведенными там таблицами.

Содержание. Элементарная вероятность $p(u)$ исхода u является *мерой реализуемости результата,* описываемого исходом u. Примеры 1—3 из § 1 поясняют это.

4.2. Вероятность P

Четвертым и самым важным элементом модели является вероятность P.

Определение 4. *Вероятностью события A называется сумма*

$$P(A) = \sum_{u \in A} p(u)$$

элементарных вероятностей $p(u)$ исходов u, составляющих событие A.

Функция P на классе событий $\mathscr{A}$, значением которой для каждого события A является его вероятность $P(A)$, называется *вероятностью.*

Пример 1. Вероятность невозможного события O равна нулю:

$$P(O)=0.$$

Пример 2. Вероятность достоверного события U равна единице:

$$P(U)=1.$$

Пример 3. В примерах 1—3 из § 1 рассматривались вероятности событий

$$P(A)=9/10,\ 3/100,\ 3/10.$$

Содержание. Вероятность $P(A)$ события A служит *мерой реализуемости явления*, описываемого событием A.

При описании каждого опыта, возможные результаты которого зависят от случая, возникает проблема выбора вероятности P, служащей удовлетворительной мерой реализуемости явлений, определяемых возможными результатами опыта. Общего правила для выбора вероятности P не существует. Используется, например, равновозможность исходов. Нередко вероятность события оценивается частотой его появления. Примеры 1—3 из § 1 поясняют сказанное.

5.2. Конечная вероятностная модель (U, $\mathscr{A}$; p, P)

Конечное множество исходов U, алгебра событий $\mathscr{A}$, элементарная вероятность p и вероятность P образуют конечную вероятностную модель (U, $\mathscr{A}$; p, P).

Пример 1. Модель Лапласа.

Пример 2. Модель Бернулли.

Пример 3. Модели примеров 1—3 из § 1.

Содержание. Конечная вероятностная модель описывает *опыт* с конечным множеством возможных результатов, реализация которых зависит от случая.

6.2. Механическая интерпретация

Будем иногда для наглядности параллельно с вероятностным языком употреблять механический язык. Условимся также называть:

исход — точкой,
событие — телом,
элементарную вероятность исхода — элементарной массой точки,
вероятность события — массой тела,
вероятностную модель — механической системой.

1.6.2. Пример 1

Модель Лапласа одного подбрасывания игральной кости представляет также механическую систему из множества точек 1, 2, 3,

4, 5, 6 на прямой, в каждой из которых находится элементарная масса, равная 1/6 (рис. 14).

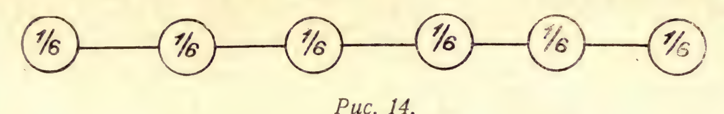


Рис. 14.

2.6.2. Пример 2

Модель Бернулли двух подбрасываний симметричной монеты представляет также механическую систему из множества точек (0, 0), (1, 0), (0, 1), (1, 1) на плоскости, в каждой из которых находится элементарная масса, равная 1/4 (рис. 15).

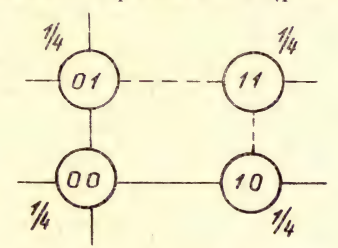


Рис. 15.

3.6.2. Пример 3

Модель примера 1 из § 1 одного выстрела по мишени представляет также механическую систему из множества точек $u=0$, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 на прямой, в которых находятся элементарные массы $p(u)=1/100$, 1/100, 1/100, 1/100, 1/100, 5/100, 5/100, 5/100, 10/100, 20/100, 50/100 (рис. 16).

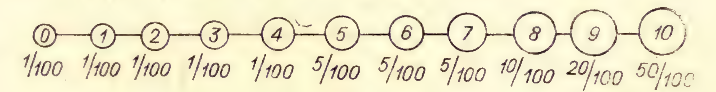


Рис. 16.

§ 3. СВОЙСТВА ВЕРОЯТНОСТИ

Вероятность P является положительной, нормированной и аддитивной функцией.

1.3. Основные свойства

Рассмотрим конечную вероятностную модель $(U, \mathscr{A}; p, P)$.

1-е основное свойство: *вероятность P является положительной функцией на классе событий $\mathscr{A}$.*

Доказательство. Положительность вероятности P следует из положительности элементарной вероятности p:

$$P(A) = \sum_{u \in A} p(u) \geqslant 0$$

для каждого события A.

2-е основное свойство: *вероятность P является нормированной функцией на классе событий $\mathscr{A}$.*

Доказательство. Нормированность вероятности P следует из нормированности элементарной вероятности p:

$$P(U) = \sum_{u \in U} p(u) = 1.$$

Прежде, чем сформулировать 3-е основное свойство, определим понятие аддитивности.

Условимся говорить, что для функции P на классе событий $\mathscr{A}$ верно правило сложения, если значение $P(A+B)$ функции P для суммы $A+B$ каждых непересекающихся событий A и B равно сумме $P(A)+P(B)$ значений $P(A)$ и $P(B)$ функции P для событий A и B:

Правило сложения:

$$P(A+B) = P(A) + P(B) \quad (AB = 0).$$

Определение. *Функция P на классе событий $\mathscr{A}$ называется аддитивной, если для нее верно правило сложения.*

Пример 1. Вероятность P в модели Лапласа является аддитивной функцией.

Пример 2. Вероятность P в модели Бернулли является аддитивной функцией.

3-е основное свойство: *вероятность P является аддитивной функцией на классе событий $\mathscr{A}$.*

Доказательство. Аддитивность вероятности P следует из ассоциативности сложения. Рассмотрим произвольные непересекающиеся события A и B.

Сумма элементарных вероятностей $p(u)$ исходов u события $A+B$ равна сумме элементарных вероятностей $p(u)$ исходов u события A плюс сумме элементарных вероятностей $p(u)$ исходов u события B. Следовательно,

$$P(A+B) = \sum_{u \in A+B} p(u) = \sum_{u \in A} p(u) + \sum_{u \in B} p(u) = P(A) + p(B)$$

для каждых непересекающихся событий A и B.

Пример. Если $A = \{1, 3\}$ и $B = \{2, 4\}$, то $A+B = \{1, 2, 3, 4\}$ и

$$P(A+B) = p(1) + p(2) + p(3) + p(4) = [p(1) + p(3)] + [p(2) + p(4)] = P(A) + P(B).$$

Замечание. В механической интерпретации правило сложения означает, что общая масса двух непересекающихся тел равна сумме масс этих тел.

Три основные свойства вероятности выражает объединяющая их

Теорема. *Вероятность P является положительной, нормированной и аддитивной функцией на классе событий $\mathcal{A}$.*

Все общие свойства вероятности являются следствиями ее положительности, нормированности и аддитивности. Предложения, выражающие эти общие свойства,— следствия доказанной теоремы.

2.3. Следствия

Сформулируем несколько общих свойств вероятности P. Рассмотрим произвольные события A и B.

1. Правило вычитания:

$$P(B-A)=P(B)-P(A) \quad (A \subseteq B).$$

Доказательство. Если $A \subseteq B$, то

$$B=A+(B-A), \quad A(B-A)=0,$$

откуда вследствие аддитивности вероятности

$$P(B)=P(A)+P(B-A), \quad P(B-A)=P(B)-P(A).$$

Доказательство правила вычитания поясняет рис. 17.

Замечание. В механической интерпретации правило вычитания означает, что если от тела B отнять часть A, то масса оставшейся части $B-A$ будет равна разности масс тел B и A.

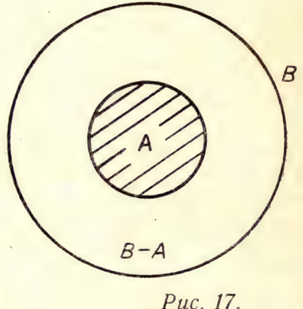


Рис. 17.

Задача 1. *Какова вероятность того, что при выстреле по мишени получается число очков большее трех, но не большее шести?*

Решение. Для решения этой задачи используем модель примера 1 из § 1. Она сводится к вычислению вероятности события $B-A$ — разности событий $B=\{3, 4, 5, 6, 7, 8, 9, 10\}$ и $A=\{6, 7, 8, 9, 10\}$. По таблице примера 1 из § 1 находим:

$$P(A)=90/100, \quad P(B)=97/100.$$

Отсюда по правилу вычитания получаем:

$$P(B-A)=P(B)-P(A)=7/100.$$

Этот результат получается и при непосредственном вычислении вероятности события $B-A=\{3, 4, 5\}$:

$$P(B-A)=p(3)+p(4)+p(5)=1/100+1/100+5/100=7/100.$$

2. Правило дополнения:

$$P(A')=1-P(A).$$

Доказательство. Используя правило вычитания и нормированность вероятности, получаем:

$$P(A')=P(U-A)=P(U)-P(A)=1-P(A).$$

Задача 2. *Какова вероятность того, что выбранная наугад из книги буква не является гласной?*

Решение. Для решения этой задачи используем модель примера 2 из § 1. Она сводится к вычислению вероятности дополнения события $A=\{$а, е, и, о, у, ы, э, ю, я$\}$. Употребляя правило дополнения и результат примера 2 из § 1, получаем:

$$P(A')=1-P(A)=1-0{,}350=0{,}750.$$

3. Правило объединения:

$$P(A\cup B)=P(A)+P(B)-P(AB)$$

Доказательство. Заметим, что

$$A\cup B=A+(B-AB),\quad A(B-AB)=0.$$

Отсюда по правилам сложения и вычитания получаем:

$$P(A\cup B)=P[A+(B-AB)]=P(A)+P(B-AB)=$$
$$=P(A)+P(B)-P(AB).$$

Замечание. В механической интерпретации правило объединения означает, что общая масса тел A и B равна сумме массы тела A и массы тела B без массы пересечения AB этих тел (рис. 18).

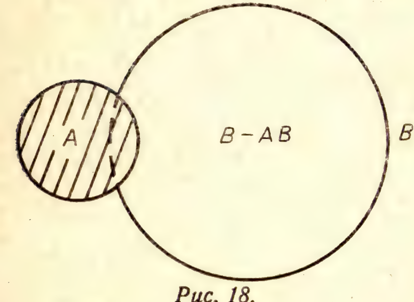


Рис. 18.

Разобьем множество натуральных чисел $\{1, \ldots, 100\}$ на десятки $\{1, \ldots, 10\}$, $\{11, \ldots, 20\}, \ldots$ $\ldots, \{91, \ldots, 100\}$.

Задача 3. *Какова вероятность того, что в выбранном наугад десятке оказывается нечетное или строго большее двух количество простых чисел?*

Решение. Для решения этой задачи используем вторую модель примера 3 из § 1. Она сводится к вычислению вероятности объединения $A\cup B$ событий $A=\{1, 3\}$ и $B=\{3, 4\}$.

По правилу объединения получаем:

$$P(A\cup B)=P(A)+P(B)-P(AB)=3/10+4/10-2/10=5/10.$$

Из правила объединения и положительности вероятности вытекает

Следствие. $P(A\cup B)\leqslant P(A)+P(B)$.

4. Правило неравенства:

$$P(A)\leqslant P(B)\quad (A\subseteq B).$$

Доказательство. Используя условие $A \subseteq B$, правило вычитания и положительность вероятности P, получаем:

$$P(B) - P(A) = P(B - A) \geqslant 0.$$

Замечание. В механической интерпретации правило неравенства означает, что при увеличении тела возрастает его масса.

Пример. Рассмотрим события $A = \{6, 7, 8, 9, 10\}$ и $B = \{3, 4, 5, 6, 7, 8, 9, 10\}$ из задачи 2. Имеем:

$$A \subseteq B, \quad P(A) = 90/100 \leqslant 97/100 = P(B).$$

Из положительности, нормированности и правила неравенства для вероятности P вытекает

Следствие. $0 \leqslant P(A) \leqslant 1$.

Замечание. При этом

$$P(O) = 0, \quad P(U) = 1.$$

Правило дополнения устанавливает эквивалентность этих равенств:

$$P(O) = P(U') = 1 - P(U) = 1 - 1 = 0,$$
$$P(U) = P(O') = 1 - P(O) = 1 - 0 = 1.$$

3.3. Общее правило объединения

При вычислении вероятностей различных событий часто бывает полезно общее выражение для вероятности объединения конечного семейства событий.

Рассмотрим произвольное натуральное число $n \geqslant 1$ и семейство $(A_i)_{i \in I}$ событий A_i, множество индексов I которого составлено из n элементов. Для каждого номера $m = 1, \ldots, n$ рассмотрим класс $\mathscr{P}_m$ всех частей множества I, составленных из m индексов. Для каждой такой части K определены пересечение $\bigcap_{i \in K} A_i$ событий A_i с индексами $i \in K$, вероятность $P\left(\bigcap_{i \in K} A_i\right)$ этого пересечения и сумма $S_m = \sum_{K \in \mathscr{P}_m} P\left(\bigcap_{i \in K} A_i\right)$ вероятностей всех таких пересечений.

Пример.

$I = \{1, 2\}$;

$$\mathscr{P}_1 = \{\{1\}, \{2\}\}, \quad \mathscr{P}_2 = \{\{1, 2\}\};$$

$$\bigcap_{i \in \{1\}} A_i = A_1, \quad \bigcap_{i \in \{2\}} A_i = A_2, \quad \bigcap_{i \in \{1,2\}} A_i = A_1 A_2;$$

$$P\left(\bigcap_{i \in \{1\}} A_i\right) = P(A_1), \quad P\left(\bigcap_{i \in \{2\}} A_i\right) = P(A_2),$$

$$P\left(\bigcap_{i \in \{1,2\}} A_i\right) = P(A_1 A_2);$$

$$S_1 = P(A_1) + P(A_2), \quad S_2 = P(A_1 A_2).$$

Используя данные обозначения, можно записать правило объ-

единения для произвольных событий A_1 и A_2 следующим образом:

$$P(A_1 \cup A_2) = S_1 - S_2.$$

Нетрудно проверить, что для каждых событий A_1, A_2, A_3 верно равенство

$$P(A_1 \cup A_2 \cup A_3) = S_1 - S_2 + S_3.$$

В самом деле, используя правило объединения для событий $A = A_1 \cup A_2$ и $B = A_3$, **получаем:**

$$P(A_1 \cup A_2 \cup A_3) = P((A_1 \cup A_2) \cup A_3) = P(A_1 \cup A_2) + P(A_3) -$$
$$- P((A_1 \cup A_2) A_3) = P(A_1 \cup A_2) + P(A_3) - P(A_1A_3 \cup A_2A_3).$$

Используя правило объединения для событий $A = A_1$, $B = A_2$ и $A = A_1A_3$, $B = A_2A_3$, имеем:

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) - P(A_1A_2) + P(A_3) -$$
$$- P(A_1A_3) - P(A_2A_3) + P(A_1A_2A_3) = S_1 - S_2 + S_3.$$

Используя принцип индукции, легко доказать

Общее правило объединения:

$$P\left(\bigcup_{i \in I} A_i\right) = \sum_{1 \leqslant m \leqslant n} \left[(-1)^{m-1} \sum_{K \in \mathcal{P}_m} P\left(\bigcap_{i \in K} A_i\right)\right].$$

Доказательство. Рассмотрим множество M всех номеров m таких, что общее правило объединения верно для каждого семейства $(A_i)_{i \in I}$ событий A_i, множество индексов I которого составлено из $n = m + 1$ элементов.

1. Номер $0 \in M$, так как для каждого семейства $(A_i)_{i \in I}$, множество индексов I которого составлено из $n = 0 + 1 = 1$ элемента i, общее правило объединения эквивалентно равенству $P(A_i) = P(A_i)$.

2. Для каждого номера l верно предложение: если $l \in M$, то $l + 1 \in M$. Действительно, рассмотрим произвольное семейство $(A_i)_{i \in \bar{I}}$ событий A_i, множество индексов $\bar{I}$ которого составлено из $\bar{n} = (l + 1) + 1$ элементов. Для каждого индекса $i_0 \in \bar{I}$ определены семейства $(A_i)_{i \in I}$ и $(A_iA_{i_0})_{i \in I}$ событий A_i и $A_iA_{i_0}$, множество индексов $I = \bar{I} - \{i_0\}$ которых составлено из $n = l + 1$ элементов. Используя правило объединения для событий $A = \bigcup_{i \in I} A_i$ и $B = A_{i_0}$, получаем:

$$(*) \quad P\left(\bigcup_{i \in \bar{I}} A_i\right) = P\left(\bigcup_{i \in I} A_i\right) + P(A_{i_0}) - P\left(\bigcup_{i \in I} A_iA_{i_0}\right).$$

Если $l \in M$, то

$$(**) \quad P\left(\bigcup_{i \in I} A_i\right) = \sum_{1 \leqslant m \leqslant n} (-1)^{m-1} \sum_{K \in \mathcal{P}_m} P\left(\bigcap_{i \in K} A_i\right),$$

$$P\left(\bigcup_{i \in I} A_iA_{i_0}\right) = \sum_{1 \leqslant m \leqslant n} (-1)^{m-1} \sum_{K \in \mathcal{P}_m} P\left(\bigcap_{i \in K} A_iA_{i_0}\right).$$

Для каждого номера $\bar m=1, \ldots, \bar n$ рассмотрим класс $\bar{\mathcal{P}}_{\bar m}$ всех частей $\bar K$ множества $\bar I$, составленных из $\bar m$ элементов. Если индекс i_0 не принадлежит множеству $\bar K$, то это множество является частью множества I и принадлежит классу $\mathcal{P}_{\bar m}$. Если индекс i_0 принадлежит множеству $\bar K$, то множество $K=\bar K-\{i_0\}$ является частью множества I и принадлежит классу $\mathcal{P}_m$ с номером $m=\bar m-1$. В соответствии с этим

$$(***)\quad \sum_{\bar K\in\bar{\mathcal{P}}_{\bar m}} P\Big(\bigcap_{i\in\bar K} A_i\Big)=\sum_{K\in\mathcal{P}_{\bar m}} P\Big(\bigcap_{i\in K} A_i\Big)+\sum_{K\in\mathcal{P}_m} P\Big(\bigcap_{i\in K} A_iA_{i_0}\Big).$$

В частности, если $\bar m=1$, то $m=0$, $\mathcal{P}_{\bar m}=\{\{i\}: i\in I\}$, $\mathcal{P}_m=0$ и

$$\sum_{\bar K\in\bar{\mathcal{P}}_{\bar m}} P\Big(\bigcap_{i\in\bar K} A_i\Big)=P(A_{i_0})+\sum_{i\in I} P(A_i).$$

Если $\bar m=\bar n$, то $m=n$, $\mathcal{P}_{\bar m}=0$, $\mathcal{P}_m=\{I\}$ и

$$\sum_{\bar K\in\bar{\mathcal{P}}_{\bar m}} P\Big(\bigcap_{i\in\bar K} A_i\Big)=P\Big(\bigcap_{i\in I} A_iA_{i_0}\Big).$$

Подставим правые части равенств (**) в правую часть равенства (*). Выделим слагаемое с номером $m=1$ в первой сумме и слагаемое с номером $m=n$ во второй сумме. Сгруппируем каждое слагаемое с номерами $m=2, \ldots, n$ первой суммы со слагаемыми с номерами $m=1, \ldots, n-1$ второй суммы. Получаем:

$$P\Big(\bigcup_{i\in\bar I} A_i\Big)=\Big(P(A_{i_0})+\sum_{i\in I} P(A_i)\Big)+\sum_{2\leqslant m\leqslant n}(-1)^{m-1}\sum_{K\in\mathcal{P}_m} P\Big(\bigcap_{i\in K} A_i\Big)+$$
$$+\sum_{1\leqslant m\leqslant n-1}(-1)^m\sum_{K\in\mathcal{P}_m} P\Big(\bigcap_{i\in K} A_iA_{i_0}\Big)+(-1)^nP\Big(\bigcap_{i\in I} A_iA_{i_0}\Big)=$$
$$=\Big(P(A_{i_0})+\sum_{i\in I} P(A_i)\Big)+\sum_{2\leqslant\bar m\leqslant n}(-1)^{\bar m-1}\Big(\sum_{K\in\mathcal{P}_{\bar m}} P\Big(\bigcap_{i\in K} A_i\Big)+$$
$$+\sum_{K\in\mathcal{P}_{\bar m-1}} P\Big(\bigcap_{i\in K} A_iA_{i_0}\Big)\Big)+(-1)^nP\Big(\bigcap_{i\in I} A_iA_{i_0}\Big).$$

Отсюда, используя равенства (***), получаем:

$$P\Big(\bigcap_{i\in\bar I} A_i\Big)=\sum_{\bar K\in\bar{\mathcal{P}}_1} P\Big(\bigcap_{i\in\bar K} A_i\Big)+\sum_{2\leqslant\bar m\leqslant\bar n-1}(-1)^{\bar m-1}\sum_{\bar K\in\bar{\mathcal{P}}_{\bar m}} P\Big(\bigcap_{i\in\bar K} A_i\Big)+$$
$$+(-1)^{\bar n-1}\sum_{\bar K\in\bar{\mathcal{P}}_{\bar n}} P\Big(\bigcap_{i\in\bar K} A_i\Big)=\sum_{1\leqslant\bar m\leqslant\bar n}(-1)^{\bar m-1}\sum_{\bar K\in\bar{\mathcal{P}}_{\bar m}} P\Big(\bigcap_{i\in\bar K} A_i\Big).$$

Общее правило объединения верно для рассматриваемого семейства $(A_i)_{i\in\bar I}$, т. е. $l+1\in M$.

По принципу индукции из пунктов 1 и 2 следует, что $M=N$: общее правило объединения верно для каждого семейства событий $(A_i)_{i\in I}$ с конечным множеством индексов I.

Замечание. Используя сокращенные обозначения для сумм, можно записать общее правило объединения также следующим образом:

$$P\left(\bigcup_{i\in I} A_i\right) = S_1 - S_2 + \ldots + (-1)^{n-1}S_n.$$

Из общего правила объединения вытекает общее правило сложения для семейства $(A_i)_{i\in I}$ событий A_i, которые попарно не пересекаются: $A_iA_j=0$ для каждых различных индексов i и j из множества I.

Общее правило сложения:

$$P\left(\sum_{i\in I} A_i\right) = \sum_{i\in I} P(A_i) \qquad (A_iA_j = 0).$$

Доказательство. Если события A_i семейства $(A_i)_{i\in I}$ попарно не пересекаются, то $\bigcap_{i\in K} A_i = 0$, $P\left(\bigcap_{i\in K} A_i\right) = 0$ для каждой части K множества индексов I, составленной строго больше чем из одного элемента. Поэтому

$$S_m = \sum_{K\in\mathcal{P}_m} P\left(\bigcap_{i\in K} A_i\right) = 0$$

для каждого номера $m>1$.

4.3. Примеры

Рассмотрим несколько примеров, поясняющих использование общих правил сложения и объединения.

1.4.3. Пример 1. Вероятность m успехов

Рассмотрим модель Бернулли n испытаний с вероятностью успеха $a (0\leqslant a\leqslant 1)$. Для каждого номера $m=0, \ldots, n$ обозначим символом A_m событие, составленное из всех исходов $u=u_1 \ldots u_n$, сумма $s(u)=u_1+\ldots+u_n$ элементов которых равна m:

$$A_m = \{u : s(u) = m\}.$$

Можно также сказать, что множество A_m составлено из всех строк u, в которых m единиц и $n-m$ нулей. В частности, если $n=3$, то

$$A_0 = \{000\},\ A_1 = \{100,\ 010,\ 001\},\ A_2 = \{110,\ 101,\ 011\},$$
$$A_3 = \{111\}.$$

Событие A_m описывает появление m успехов и $n-m$ неудач при n испытаниях, описываемых моделью Бернулли.

Задача. *Какова вероятность того, что при n одинаковых и независимых испытаниях, каждое из которых с вероятностью a оканчивается успехом, а с вероятностью $1-a$ — неудачей, m испытаний оканчивается успехом, а $n-m$ — неудачей?*

Решение. В модели Бернулли n испытаний с вероятностью успеха a задача сводится к вычислению вероятности $P(A_m)$ события A_m. Заметим, что в этой модели элементарная вероятность каждого исхода u, принадлежащего событию A_m, равна числу $a^m(1-a)^{n-m}$:

$$p(u)=a^{s(u)}(1-a)^{n-s(u)}=a^m(1-a)^{n-m}(s(u)=m).$$

Поэтому вычисление вероятности $P(A_m)=\sum_{u\in A_m} p(u)$ события A_m сводится к вычислению числа исходов u, составляющих событие A_m. Число этих исходов равно числу выборок по m номеров мест, на которых находятся единицы, из n номеров всех мест в строке u. Число таких выборок равно $\binom{n}{m}$. Таким образом,

$$P(A_m)=\binom{n}{m}a^m(1-a)^{n-m}.$$

В частности, если $n=3$, то

$$P(A_0)=(1-a)^3,\ P(A_1)=2a(1-a)^2,\ P(A_2)=2a^2(1-a),$$
$$P(A_3)=a^3.$$

События A_m попарно не пересекаются. Используя общее правило сложения и формулу Ньютона для бинома, получаем:

$$P\left(\sum_{0\leqslant m\leqslant n} A_m\right)=\sum_{0\leqslant m\leqslant n} P(A_m)=$$
$$=\sum_{0\leqslant m\leqslant n}\binom{n}{m}a^m(1-a)^{n-m}=(a+(1-a))^n=1.$$

Например, если $n=3$, то

$$P(A_0+A_1+A_2+A_3)=P(A_0)+P(A_1)+P(A_2)+P(A_3)=$$
$$=a(1-a)^3+2a(1-a)^2+2a^2(1-a)+a^3=(a(a-a))^3=1.$$

Тот же самый результат вытекает из нормированности вероятности. Каждый исход u принадлежит одному из событий A_m:

$$\sum_{0\leqslant m\leqslant n} A_m=U.$$

Следовательно,

$$P\left(\sum_{0\leqslant m\leqslant n} A_m\right)=P(U)=1.$$

2.4.3. Пример 2. Задача о совпадениях

Классическую задачу о хотя бы одном совпадении можно сформулировать следующим образом.

Задача 1. *Какова вероятность того, что при тщательном перемешивании колоды хотя бы одна карта останется на своем месте?*

Решение. Обозначим число карт в колоде буквой n. Припишем каждой карте номер ее места в рассматриваемой колоде до тасования. Возможные результаты перемешивания удобно описывать перестановками множества $I=\{1, \ldots, n\}$ номеров карт. Число таких перестановок равно $n!$. Тщательность перемешивания означает, что все эти перестановки равновозможны. Таким образом, для описания рассматриваемого опыта можно использовать модель Лапласа для $n!$ равновозможных исходов. В качестве множества исходов U удобно взять множество всех перестановок u номеров $1, \ldots, n$.

Для каждого номера $i=1, \ldots, n$ обозначим символом A_i событие, образованное всеми перестановками u, оставляющими на месте номер i:

$$A_i=\{u : u(i)=i\}.$$

Объединение

$$A=\bigcup_{1 \leqslant i \leqslant n} A_i$$

событий A_i образовано всеми перестановками u, оставляющими на месте хотя бы один номер из множества $I=\{1, \ldots, n\}$.

В частности, если $n=3$, то

$$A_1=\left\{\begin{pmatrix}1\,2\,3\\1\,2\,3\end{pmatrix}, \begin{pmatrix}1\,2\,3\\1\,3\,2\end{pmatrix}\right\}, A_2=\left\{\begin{pmatrix}1\,2\,3\\1\,2\,3\end{pmatrix}, \begin{pmatrix}1\,2\,3\\3\,2\,1\end{pmatrix}\right\}, A_3=\left\{\begin{pmatrix}1\,2\,3\\1\,2\,3\end{pmatrix}, \begin{pmatrix}123\\213\end{pmatrix}\right\}$$

$$A=A_1 \cup A_2 \cup A_3=\left\{\begin{pmatrix}1\,2\,3\\1\,2\,3\end{pmatrix}, \begin{pmatrix}1\,2\,3\\1\,3\,2\end{pmatrix}, \begin{pmatrix}1\,2\,3\\3\,2\,1\end{pmatrix}, \begin{pmatrix}1\,2\,3\\2\,1\,3\end{pmatrix}\right\}.$$

Решение задачи о совпадении сводится к вычислению вероятности $P(A)$ события A.

Используем для вычисления вероятности $P(A)$ события A общее правило объединения. Рассмотрим произвольную часть K множества $I=\{1, \ldots, n\}$, составленную из m номеров. Число элементов пересечения $\bigcap_{i \in K} A_i$ равно числу перестановок множества $I-K$, составленного из $n-m$ номеров. Число таких перестановок равно $(n-m)!$. Следовательно, для каждой такой части K

$$P\left(\bigcap_{i \in K} A_i\right)=\frac{(n-m)!}{n!}.$$

Число выборок K по m из n номеров равно $\binom{n}{m}$. Поэтому

$$S_m=\sum_{K \in \mathscr{P}_m} P\left(\bigcap_{i \in K} A_i\right)=\binom{n}{m} \frac{(n-m)!}{n!}=\frac{n!\,(n-m)!}{m!\,(n-m)!\,n!}=\frac{1}{m!}.$$

Используя общее правило объединения, получаем:

$$P(A)=\frac{1}{1!}-\frac{1}{2!}+\frac{1}{3!}-\frac{1}{4!}+\ldots+(-1)^{n-1}\frac{1}{n!}.$$

Для $n=1, 2, 3, 4, 5, 6, 7, 8$ вероятности $P(A)$ выражает следующая таблица:

n	1	2	3	4	5	6	7	8
$P(A)$	1	0,50000	0,66667	0,62500	0,63333	0,63196	0,63214	0,63212

Замечание. Можно доказать, что для $n>8$ карт в колоде вероятность $P(A)$ того, что при тщательном перемешивании хотя бы одна карта остается на своем месте, практически постоянна и приближенно равна 0,63212. Точнее,

$$P(A)\approx 1-1/e,$$

где буква e обозначает знаменитое число Эйлера — 2,7182818...

Поучительно попытаться объяснить качественными рассуждениями полученный результат: для достаточно больших колод вероятность $P(A)$ не мала и не велика и практически не зависит от числа карт в колоде.

Интересно также провести серию опытов с тщательным перемешиванием обычной колоды карт и сравнить с вероятностью $P(A)$ наблюденную частоту случаев, когда хотя бы одна карта осталась на своем месте.

Решение задачи 1 позволяет легко решить задачу об отсутствии совпадений.

Задача 2. *Какова вероятность того, что при тщательном перемешивании колоды ни одна карта не останется на своем месте?*

Решение. Используем ту же модель, что и при решении задачи 1. Она сводится к вычислению вероятности $P(A')$ события A', дополнительного событию A. Событие A' состоит из всех перестановок u, не оставляющих на месте ни одного номера из множества $I=\{1, \ldots, n\}$. В частности, если $n=3$, то

$$A'=\left\{\begin{pmatrix}1\,2\,3\\3\,1\,2\end{pmatrix}, \begin{pmatrix}1\,2\,3\\2\,3\,1\end{pmatrix}\right\}.$$

По правилу дополнения

$$P(A')=1-P(A)=1-\frac{1}{1!}+\frac{1}{2!}-\frac{1}{3!}+\ldots+(-1)^{n}\frac{1}{n!}.$$

Для $n=1, 2, 3, 4, 5, 6, 7, 8$ вероятности $P(A')$ можно получить, используя таблицу для вероятностей $P(A)$. Для $n>8$ карт в колоде вероятность $P(A')$ того, что при тщательном перемешивании ни одна карта не останется на своем месте, практически постоянна и приближенно равна 0,36788. Точнее,

$$P(A')=1/e.$$

Решение задачи 2 позволяет легко решить задачу о данном числе совпадений.

Задача 3. *Какова вероятность того, что при тщательном перемешивании колоды из n карт ровно m останутся на своих местах?*

Решение. Используем ту же модель, что и при решении задач 1 и 2. Для каждой части K множества $I=\{1, \ldots, n\}$, составленной из $m \neq n-1$ номеров, обозначим символом B_K событие, составленное из всех перестановок u, оставляющих на месте каждый номер i из множества K и переставляющих на другое место каждый номер j не из множества K:

$$B_K = \{u : u(i) = i,\ i \in K,\ u(j) \neq j,\ j \notin K\}.$$

В частности, если $n=3$, то

$$B_0 = \left\{\begin{pmatrix}1\,2\,3\\3\,1\,2\end{pmatrix}, \quad \begin{pmatrix}1\,2\,3\\2\,3\,1\end{pmatrix}\right\};$$

$$B_{\{1\}} = \left\{\begin{pmatrix}1\,2\,3\\1\,3\,2\end{pmatrix}\right\}, \quad B_{\{2\}} = \left\{\begin{pmatrix}1\,2\,3\\3\,2\,1\end{pmatrix}\right\}, \quad B_{\{3\}} = \left\{\begin{pmatrix}1\,2\,3\\2\,1\,3\end{pmatrix}\right\};$$

$$B_I = \left\{\begin{pmatrix}1\,2\,3\\1\,2\,3\end{pmatrix}\right\}.$$

Рассмотрим семейство $(B_k)_{K \in \mathcal{P}_m}$ таких событий B_K. По предположению $m \neq n-1$. В этом случае события B_K попарно не пересекаются. Сумма

$$B = \sum_{K \in \mathcal{P}_m} B_K$$

событий B_K состоит из всех перестановок u, оставляющих на месте ровно m номеров. Решение общей задачи о совпадении сводится к вычислению вероятности $P(B)$ события B.

Используем для вычисления вероятности $P(B)$ события B общее правило сложения. Вычислим предварительно вероятность $P(B_K)$ события B_K для каждого множества $K \in \mathcal{P}_m$.

Число элементов множества B_K равно числу перестановок множества $I-K$, не оставляющих на месте ни одного номера. Используя решение задачи 2, получаем:

$$n(B_K) = (n-m)! \sum_{0 \leqslant l \leqslant n-m} (-1)^l \frac{1}{l!},$$

поэтому

$$P(B_K) = \frac{(n-m)!}{n!} \sum_{0 \leqslant l \leqslant n-m} (-1)^l \frac{1}{l!}$$

для каждого множества $K \in \mathscr{P}_m$. Так как всего таких множеств $\binom{n}{m}$, то из общего правила сложения следует, что

$$P(B) = \sum_{K \in \mathscr{P}_m} P(B_K) = \binom{n}{m} \frac{(n-m)!}{n!} \sum_{0 \leqslant l \leqslant n-m} (-1)^l \frac{1}{l!} =$$

$$= \frac{n!\,(n-m)!}{m!\,(n-m)!\,n!} \sum_{0 \leqslant l \leqslant n-m} (-1)^l \frac{1}{l!} = \frac{1}{m!} \sum_{0 \leqslant l \leqslant n-m} (-1)^l \frac{1}{l!}.$$

В частности, для $m = n$ имеем:

$$P(B) = \frac{1}{n!}.$$

Значит, оставляет на месте все номера множества $I = \{1, \ldots, n\}$ одна единственная тождественная перестановка.

Замечание. Равенство для вероятности $P(B)$ было получено в предположении, что $m \neq n-1$. Однако оно верно и при $m = n-1$. В этом случае

$$P(B) = \frac{1}{(n-1)!}\left(\frac{1}{0!} - \frac{1}{1!}\right) = 0.$$

Это равенство соответствует тому факту, что невозможно оставить $n-1$ номеров из n на своих местах, а один номер переставить на другое место: все другие места заняты своими номерами. В частности, если $n = 3$ и перестановка u множества $I = \{1, 2, 3\}$ оставляет на своих местах номера 1 и 2, то эта перестановка оставляет на месте и номер 3:

$$u = \begin{pmatrix} 1\,2\,3 \\ 1\,2\,\cdot \end{pmatrix} = \begin{pmatrix} 1\,2\,3 \\ 1\,2\,3 \end{pmatrix}.$$

§ 4. УСЛОВНАЯ ВЕРОЯТНОСТЬ

При описании опытов с конечным множеством возможных результатов часто приходится учитывать различные дополнительные условия, связанные с реализацией некоторых явлений. Поэтому для описывающей опыт вероятностной модели и для каждого возможного условия с помощью условной элементарной вероятности определяется условная вероятностная модель.

Рассмотрим произвольную конечную вероятностную модель $(U, \mathscr{A}; p, P)$. Условимся каждое событие, вероятность которого не равна нулю, называть возможным событием. Обозначим буквой B произвольное возможное событие для рассматриваемой модели:

$$P(B) > 0.$$

1.4. Условная элементарная вероятность p_B

Условную вероятностную модель для конечной вероятностной модели (U, $\mathcal{A}$; p, P) и возможного события B определяет условная элементарная вероятность p_B.

Определение 5. *Условной элементарной вероятностью при условии B называется функция p_B на множестве исходов U, значение $p_B(u)$ которой для каждого исхода u определяется равенствами*:

$$p_B(u) = p(u)/P(B) \quad (u \in B), \; p_B(u) = 0 \quad (u \notin B).$$

Условная элементарная вероятность p_B является положительной нормированной функцией на множестве U. Положительность p_B следует из определения и положительности элементарной вероятности p, а нормированность p_B вытекает из равенств

$$\sum_{u \in U} p_B(u) = \sum_{u \in B} p_B(u) = \sum_{u \in B} p(u)/P(B) = P(B)/P(B) = 1.$$

Примеры. Примеры 1—3 из § 4 главы 1.

Содержание. Условная элементарная вероятность p_B является мерой реализуемости результатов описываемого моделью (U, $\mathcal{A}$; p, P) опыта при дополнительном условии, что реализуется явление, описываемое событием B.

Для каждого результата, описываемого исходом из B, мера реализуемости выбрана пропорциональной элементарной вероятности этого исхода. В качестве коэффициента пропорциональности взято число $1/P(B)$. Для каждого результата, описываемого исходом не из B, мера реализуемости выбрана равной нулю. Выбор коэффициента пропорциональности обеспечивает нормированность условной элементарной вероятности p_B.

2.4. Условная вероятность P_B

Условная элементарная вероятность p_B определяет условную вероятность P_B так же, как элементарная вероятность p определяет вероятность P.

Определение 6. *Условной вероятностью при условии B называется функция P_B на классе событий $\mathcal{A}$, значение $P_B(A)$ которой для каждого события A определяется равенством*

$$P_B(A) = \sum_{u \in A} p_B(u).$$

Примеры. Примеры 1—3 из § 4 главы 1.

Содержание. Условная вероятность P_B служит мерой реализуемости явлений, определяемых результатами описываемого моделью (U, $\mathcal{A}$; p, P) опыта при дополнительном условии, что реализуется явление, описываемое событием B.

Для явления, описываемого событием A, эта мера пропорциональна мере совместной реализуемости явлений, описываемых со-

бытиями A и B. Коэффициентом пропорциональности является число $1/P(B)$, что обеспечивает нормированность условной вероятности P_B.

3.4. Условная вероятностная модель $(U, \mathscr{A}; p, P)$

Конечная вероятностная модель $(U, \mathscr{A}; p, P)$ и возможное событие B для этой модели определяют конечную вероятностную модель $(U, \mathscr{A}; p_B, P_B)$. Множество исходов U и алгебра событий $\mathscr{A}$ в этой условной модели те же, что и в исходной модели. Условные элементарную вероятность p_B и вероятность P_B дают определения 5 и 6.

Замечание. В механической интерпретации переход от модели $(U, \mathscr{A}; p, P)$ к условной модели $(U, \mathscr{A}; p_B, P_B)$ означает перенос всей массы тела U на тело B, при котором элементарная масса каждой точки тела B пропорционально увеличивается. При этом масса каждого тела становится пропорциональной массе пересечения тел A и B.

В частности, для модели Лапласа одного подбрасывания игральной кости с множеством исходов $U=\{1, 2, 3, 4, 5, 6\}$ и события $B=\{1, 3, 5\}$, описывающего появление нечетного числа очков, этот переход на механическом языке означает перенос массы, который поясняют рис. 19 и 20.

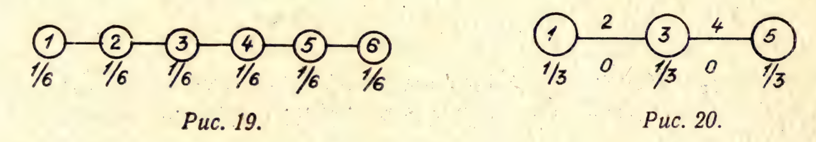


Рис. 19. *Рис. 20.*

4.4. Правила умножения и деления

Рассмотрим конечную вероятностную модель $(U, \mathscr{A}; p, P)$. Для каждых события A и возможного события B в этой модели верно

Правило деления:

$$P_B(A)=P(AB)/P(B).$$

В самом деле, из определений условной элементарной вероятности p_B, условной вероятности P_B и вероятности P следуют равенства:

$$P_B(A)=\sum_{u\in A} p_B(u)=\sum_{u\in AB} p_B(u)=\sum_{u\in AB} p(u)/P(B)=P(AB)/P(B).$$

В частности, $P_B(B')=P(B'B)/P(B)=P(O)/P(B)=0$, $P_B(B)=$ $=P(BB)/P(B)=P(B)/P(B)=1$.

Из правила деления следует

Правило умножения:

$$P(AB)=P(B)P_B(A).$$

Замечание. В механической интерпретации правило деления означает, что число $P_B(A)$ выражает долю тела A в массе тела B, правило умножения,— что масса пересечения AB тел A и B равна произведению массы тела B на долю тела A в массе тела B.

Правила умножения и деления часто используются при вычислении вероятностей событий.

Примеры. Рассмотрим несколько примеров вычисления условной вероятности.

Задача 1. *Производится один выстрел по стандартной мишени. Известно, что мишень поражается. Какова вероятность при этом условии того, что получается число очков, строго большее пяти?*

Решение. Решим эту задачу в предположении, что выстрел описывается моделью примера 1 из § 1. Для того, чтобы учесть условие о поражении мишени, рассмотрим описывающее его событие $B=\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. Событие B возможно:

$$P(B)=1-1/100=99/100.$$

Дело сводится к вычислению условной вероятности $P_B(A)$ события $A=\{6, 7, 8, 9, 10\}$. Так как $AB=A$, то, используя правило деления и решение задачи в примере 1 из § 1, получаем:

$$P_B(A)=P(AB)/P(B)=P(A)/P(B)=$$
$$=(90/100)/(99/100)=90/99.$$

Задача 2. *Из книги, написанной на русском языке, наугад выбирается буква. Известно, что выбираемая буква гласная. Какова вероятность при этом условии того, что выбирается буква а?*

Решение. Решим эту задачу в предположении, что выбор буквы описывается моделью примера 2 из § 1. Для того, чтобы учесть условие о гласности выбираемой буквы, рассмотрим описывающее его событие $B=\{$а, е, и, о, у, ы, э, ю, я$\}$. Событие B возможно: проделанные при решении задачи в примере 2 из § 1 вычисления показывают, что

$$P(B)=0{,}350.$$

Дело сводится к вычислению условной вероятности $P_B(A)$ события $A=\{a\}$. Используя правило деления и таблицу примера 2 из § 1, получаем:

$$P_B(A)=P(AB)/P(B)=p(a)/P(B)=$$
$$=0{,}062/0{,}350\approx 0{,}177.$$

Задача 3. *Наугад выбирается один из десятков первой сотни натуральных чисел. Известно, что выбирается четный десяток. Какова вероятность при этом условии того, что в выбранном десятке оказывается нечетное количество простых чисел?*

Решение. Для решения задачи используем первую модель примера 3 из § 1. Чтобы учесть условие о четности выбираемого десятка, рассмотрим описывающее это условие событие $B=\{2, 4, 6, 8, 10\}$. Событие B возможно:

$$P(B)=5/10.$$

Таблица примера 3 из § 1 показывает, что появление десятка с нечетным количеством простых чисел описывает событие $A=\{5, 8, 10\}$. Дело сводится к вычислению условной вероятности $P_B(A)$ события A. Заметим, что $AB=\{8, 10\}$ и $P(AB)=2/10$. Используя правило деления, получаем:

$$P_B(A)=P(AB)/P(B)=(2/10)/(5/10)=2/5.$$

Применение правила умножения поясняют примеры 1—3 из § 4 главы 1.

5.4. Формула полной вероятности

Рассмотрим произвольную конечную вероятностную модель $(U, \mathscr{A}; p, P)$. Каждое семейство $(B_i)_{i\in I}$ попарно не пересекающихся возможных событий B_i, сумма которых равна U, условимся называть возможным разбиением множества исходов U. Для каждого возможного разбиения $(B_i)_{i\in I}$ множества исходов U и каждого события A верна

Формула полной вероятности:

$$P(A)=\sum_{i\in I}P(B_i)P_{B_i}(A).$$

Доказательство. По условию, события B_i попарно не пересекаются и их сумма равна U. Следовательно, события AB_i попарно не пересекаются и их сумма равна A:

$$\sum_{i\in I}AB_i=A\sum_{i\in I}B_i=AU=A.$$

По общему правилу сложения отсюда следует, что

(*) $$P(A)=\sum_{i\in I}P(AB_i).$$

По предположению,

$$P(B_i)>0;$$

по правилу умножения,

(**) $$P(AB_i)=P(B_i)P_{B_i}(A)$$

для каждого индекса $i\in I$. Из равенств (*) и (**) вытекает формула полной вероятности.

Замечание. В механической интерпретации формула полной вероятности означает, что полная масса $P(A)$ тела A равна сум-

ме масс $P(AB_i)$ его частей AB_i, каждая из которых равна произведению массы $P(B_i)$ тела B_i на долю $P_{B_i}(A)$ тела A в массе тела B.

Стандартные задачи, решаемые с помощью формулы полной вероятности, можно схематически описать следующим образом.

Задача о полной вероятности. *Известны*:

1) вероятности $P(B_i)=\beta_i$ нескольких исключающих друг друга условий B_i, одно из которых с достоверностью выполняется;

2) условные вероятности $P_{B_i}(A)=\alpha_i$ события A при условии, что выполняется B_i.

Какова вероятность $P(A)$ события A?

Решение. Термины «событие» и «вероятность» в формулировке задачи используются в содержательном смысле. Построим вероятностную модель для этой задачи.

Рассмотрим конечное множество I индексов i. Для каждого индекса i рассмотрим множество

$$B_i=\{a_i, b_i\},$$

составленное из двух элементов $a_i=i1$ и $b_i=i0$. Множества B попарно не пересекаются:

$$B_iB_j=\{i1, i0\} \cap \{j1, j0\}$$

для каждых различных индексов i и j. В качестве множества исходов U для нашей модели возьмем сумму множеств B_i:

$$U=\sum_{i\in I} B_i.$$

Обозначим буквой A событие, составленное из исходов a_i:

$$A=\{a_i: i\in I\}.$$

Рассмотрим функцию p на множестве исходов U, значения $p(u)$ для каждого исхода u которой определяются равенствами:

$$p(a_i)=\beta_i\alpha_i, \quad p(b_i)=\beta_i(1-\alpha_i)$$

для каждого индекса i. Функция p положительна и нормирована. В самом деле, 1-е условие задачи предполагает, что $\beta_i>0, \sum_{i\in I}\beta_i=1$ 2-е,— что $0\leqslant\alpha_i\leqslant 1$ для каждого индекса i. Поэтому

$$p(a_j)=\beta_i\alpha_i\geqslant 0, \quad p(b_i)=\beta_i(1-\alpha_i)\geqslant 0$$

для каждого i, т. е. $p(u)\geqslant 0$ для каждого исхода u. Кроме того,

$$\sum_{u\in U} p(u)=\sum_{i\in I} p(a_i)+\sum_{i\in I} p(b_i)=\sum_{i\in I}(p(a_i)+p(b_i))=\sum_{i\in I}\beta_i=1.$$

Таким образом, рассматриваемая функция p на множестве исходов U является элементарной вероятностью. Возьмем эту функцию в качестве элементарной вероятности для нашей модели.

Вероятностная модель $(U, \mathscr{A}; p, P)$, определяемая выбранными множеством исходов U и элементарной вероятностью p, удовлетворяет условиям задачи:

1) $P(B_i) = p(a_i) + p(b_i) = \beta_i\alpha_i + \beta_i(1-\alpha_i) = \beta_i$,

2) $P_{B_i}(A) = P(AB_i)/P(B_i) = p(a_i)/P(B_i) = \beta_i\alpha_i/\beta_i = \alpha_i$

для каждого индекса i.

Решение задачи о полной вероятности сводится к применению формулы полной вероятности для построенной модели:

$$P(A) = \sum_{i\in I} P(B_i) P_{B_i}(A) = \sum_{i\in I} \beta_i\alpha_i.$$

Примеры. Использование формулы полной вероятности поясняют примеры 1—3 из § 4 главы 1. Рассмотрим еще несколько примеров.

Задача 1. *Команда составлена из двух отличных, трех хороших и пяти средних стрелков. Каждый отличный стрелок попадает в мишень с вероятностью 0,99, хороший — 0,90 и средний — 0,75. Какова вероятность того, что наугад выбранный из комады стрелок попадает в мишень?*

Решение. Используем схему задачи о полной вероятности.

Рассмотрим множество номеров $I = \{1, 2, 3\}$ и множество исходов $U = \{11, 10, 21, 20, 31, 30\}$. Исход 11 (21, 31) описывает результат «выбирается отличный (хороший, средний) стрелок, и он попадает в мишень», исход 10 (20, 30) — результат «выбирается отличный (хороший, средний) стрелок, и он не попадает в мишень».

События

$$B_1 = \{11, 10\},\ B_2 = \{21, 20\},\ B_3 = \{31, 30\}$$

описывают выбор соответственно отличного, хорошего и среднего стрелка. Они попарно не пересекаются и в сумме равны U.

Событие

$$A = \{11, 21, 31\}$$

описывает попадание в мишень выбранного стрелка.

Так как стрелок выбирается наугад из команды указанного в формулировке задачи состава, то условиям B_1, B_2, B_3 можно приписать вероятности $P(B_1) = \beta_1 = 0{,}2$; $P(B_2) = \beta_2 = 0{,}3$; $P(B_3) = \beta_3 = 0{,}5$.

Условные вероятности события A при каждом из условий B_1, B_2, B_3 указаны в формулировке задачи:

$$P_{B_1}(A) = \alpha_1 = 0{,}99;\quad P_{B_2}(A) = \alpha_2 = 0{,}90;\quad P_{B_3}(A) = \alpha_3 = 0{,}75.$$

Рассмотрим элементарную вероятность p со значениями:

$$p(11) = \beta_1\alpha_1 = 0{,}2\cdot 0{,}99;\ p(10) = \beta_1(1-\alpha_1) = 0{,}2\cdot 0{,}01;$$

$$p(21) = \beta_2\alpha_2 = 0{,}3\cdot 0{,}90;\ p(20) = \beta_2(1-\alpha_2) = 0{,}3\cdot 0{,}10;$$

$$p(31) = \beta_3\alpha_3 = 0{,}5\cdot 0{,}75;\ p(30) = \beta_3(1-\alpha_3) = 0{,}5\cdot 0{,}25.$$

Задача сводится к вычислению вероятности $P(A)$ события $A=\{11, 21, 31\}$ в модели, определяемой описанными множеством исходов U и элементарной вероятностью p на нем. По формуле полной вероятности получаем:

$$P(A)=P(B_1)P_{B_1}(A)+P(B_2)P_{B_2}(A)+P(B_3)P_{B_3}(A)=$$
$$=\beta_1\alpha_1+\beta_2\alpha_2+\beta_3\alpha_3=0,2\cdot 0,99+0,3\cdot 0,90+0,5\cdot 0,75\approx 0,82.$$

Замечание. Коротко решение задачи 1 можно изложить следующим образом. Задача 1 сводится к задаче о полной вероятности с вероятностями условий $\beta_1=0,2$; $\beta_2=0,3$; $\beta_3=0,5$ и условными вероятностями $\alpha_1=0,99$; $\alpha_2=0,90$; $\alpha_3=0,75$. По формуле полной вероятности искомая вероятность равна

$$\sum_{1\leqslant i\leqslant 3}\beta_i\alpha_i\approx 0,82.$$

Задача 2. *Имеются две книги: русская и английская. Из них наугад выбирается одна книга и из нее также наугад — одна буква. Какова вероятность того, что выбранной окажется буква «е»?*

Решение. Используем схему задачи о полной вероятности.

Рассмотрим множество номеров $I=\{1, 2\}$ и множество исходов $U=\{11, 10; 21, 20\}$. Исход 11 описывает выбор русской книги и буквы «е» из нее, исход 10 — русской книги и не буквы «е», исход 21 — выбор английской книги и буквы «е» из нее, исход 20 — английской книги и не буквы «е».

События $B_1=\{11, 10\}$, $B_2=\{21, 20\}$ описывают выбор соответственно русской и английской книги, событие $A=\{11, 21\}$ — выбор буквы «е».

Так как книга выбирается наугад из двух имеющихся, то условиям B_1, B_2 можно приписать вероятности

$$P(B_1)=\beta_1=0,5;\ P(B_2)=\beta_2=0,5.$$

В качестве условных вероятностей события A при каждом из условий B_1, B_2 возьмем частоту, с которой встречается буква «е» соответственно в русском и английском тексте:

$$P_{B_1}(A)=\alpha_1=0,072;\quad P_{B_2}(A)=\alpha_2=0,131.$$

Рассмотрим элементарную вероятность p со значениями:

$$p(11)=\beta_1\alpha_1=0,5\cdot 0,072;\ p(10)=\beta_1(1-\alpha_1)=0,5\cdot 0,928;$$
$$p(21)=\beta_2\alpha_2=0,5\cdot 0,131;\ p(20)=\beta_2(1-\alpha_2)=0,5\cdot 0,869.$$

Задача сводится к вычислению вероятности $P(A)$ события $A=\{11, 21\}$ в модели, определяемой описанными множеством исходов U и элементарной вероятностью p на нем. По формуле полной вероятности получаем:

$$P(A)=P(B_1)P_{B_1}(A)+P(B_2)P_{B_2}(A)=\beta_1\alpha_1+\beta_2\alpha_2=$$
$$=0,5\cdot 0,072+0,5\cdot 0,131+0,5\cdot 0,203\approx 0,101.$$

Замечание. Коротко решение задачи 2 можно изложить следующим образом. Задача 2 сводится к задаче о полной вероятности с вероятностями условий $\beta_1=0{,}5$; $\beta_2=0{,}5$ и условными вероятностями $\alpha_1=0{,}072$; $\alpha_2=0{,}131$. По формуле полной вероятности искомая вероятность равна

$$\sum_{1\leqslant i\leqslant 2}\beta_i\alpha_i\approx 0{,}101.$$

Задача 3. *Множество натуральных чисел $\{1, \ldots, 100\}$ разбито на десятки $\{1, \ldots, 10\}, \ldots, \{91, \ldots, 100\}$. Из них наугад выбирается один десяток и из него — одно число. Какова вероятность того, что выбранное число оказывается простым?*

Решение. Используем схему задачи о полной вероятности и пример 3 из § 1.

Рассмотрим множество номеров $I=\{1, 2, 3, 4\}$ и множество исходов $U=\{11, 10;\ 21, 20;\ 31, 30;\ 41, 40\}$. Исходы 11, 21, 31, 41 описывают выбор десятка соответственно с 1, 2, 3, 4 простыми числами и выбор из этого десятка простого числа. Исходы 10, 20, 30, 40 — выбор десятка соответственно с 1, 2, 3, 4 простыми числами и выбор из этого десятка не простого числа.

События

$$B_1=\{11, 10\},\ B_2=\{21, 20\},\ B_3=\{31, 30\},\ B_4=\{41, 40\}$$

описывают выбор десятка соответственно с 1, 2, 3, 4 простыми числами, событие $A=\{11, 21, 31, 41\}$ — выбор простого числа.

Так как десяток выбирается наугад, то таблица примера 3 из § 1 показывает, что условиям B_1, B_2, B_3, B_4 можно приписать вероятности $\beta_1=0{,}1$; $\beta_2=0{,}5$; $\beta_3=0{,}2$; $\beta_4=0{,}2$.

Условные вероятности события A при каждом из условий B_1, B_2, B_3, B_4 равны вероятностям выбрать наугад простое число из десятка чисел, среди которых соответственно 1, 2, 3, 4 простых:

$$\alpha_1=0{,}1;\ \alpha_2=0{,}2;\ \alpha_3=0{,}3;\ \alpha_4=0{,}4.$$

Рассмотрим элементарную вероятность p со значениями:

$$p(11)=\beta_1\alpha_1=0{,}1\cdot 0{,}1;\ p(10)=\beta_1(1-\alpha_1)=0{,}1\cdot 0{,}9;$$
$$p(21)=\beta_2\alpha_2=0{,}5\cdot 0{,}2;\ p(20)=\beta_2(1-\alpha_2)=0{,}5\cdot 0{,}8;$$
$$p(31)=\beta_3\alpha_3=0{,}2\cdot 0{,}3;\ p(30)=\beta_3(1-\alpha_3)=0{,}2\cdot 0{,}7;$$
$$p(41)=\beta_4\alpha_4=0{,}2\cdot 0{,}4;\ p(40)=\beta_4(1-\alpha_4)=0{,}2\cdot 0{,}6.$$

Задача сводится к вычислению вероятности $P(A)$ события $A=\{11, 21, 31, 41\}$ в модели, определяемой описанными множеством исходов U и элементарной вероятностью p на нем. По формуле полной вероятности получаем:

$$P(A)=P(B_1)P_{B_1}(A)+P(B_2)P_{B_2}(A)+P(B_3)P_{B_3}(A)+$$
$$+P(B_4)P_{B_4}(A)=\beta_1\alpha_1+\beta_2\alpha_2+\beta_3\alpha_3+\beta_4\alpha_4=$$
$$=0{,}1\cdot 0{,}1+0{,}5\cdot 0{,}2+0{,}2\cdot 0{,}3+0{,}2\cdot 0{,}4=0{,}25.$$

Замечание. Коротко решение задачи 3 можно изложить следующим образом. Задача 3 сводится к задаче о полной вероятности с вероятностями условий $\beta_1=0{,}1$; $\beta_2=0{,}5$; $\beta_3=0{,}2$; $\beta_4=0{,}2$ и условными вероятностями $\alpha_1=0{,}1$; $\alpha_2=0{,}2$; $\alpha_3=0{,}3$; $\alpha_4=0{,}4$. По формуле полной вероятности искомая вероятность равна

$$\sum_{1\leqslant i\leqslant 4}\beta_i\alpha_i=0{,}25.$$

В связи с задачей 3 естественно возникает

Задача 3′. *Из множества натуральных чисел $\{1, \ldots, 100\}$ наугад выбирается одно число. Какова вероятность того, что выбираемое число оказывается простым?*

Решение. Для решения этой задачи используем стандартную модель Лапласа $n=100$ равновероятностных исходов. Задача сводится к вычислению вероятности события

$$A=\{2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97\},$$

составленного из всех простых чисел, меньших 100. Так как таких чисел 25, то

$$P(A)=0{,}25.$$

Как и следовало ожидать, вероятность того, что в наугад выбранном из первой сотни десятке наугад выбранное число оказывается простым, равна вероятности того, что наугад выбранное из этой сотни число оказывается простым.

6.4. Формула Байеса

Рассмотрим произвольную конечную вероятностную модель $(U, \mathscr{A}; p, P)$. Для каждого возможного разбиения $(B_i)_{i\in I}$ множество исходов U и каждого возможного события A верна

Формула Байеса:

$$P_A(B_i)=P(B_i)P_{B_i}(A)\Big/\sum_{j\in I}P(B_j)P_{B_j}(A).$$

Доказательство. По предположению, событие A возможно. Следовательно, по правилу деления,

$$(*)\qquad P_A(B_i)=P(AB_i)/P(A)$$

для каждого индекса $i\in I$. Так как событие B_i возможно, то, по правилу умножения,

$$(**)\qquad P(AB_i)=P(B_i)P_{B_i}(A)$$

для каждого индекса $i\in I$. Наконец, для семейства событий $(B_j)_{j\in I}$ и события A верна формула полной вероятности:

$$(***)\qquad P(A)=\sum_{j\in I}P(B_j)P_{B_j}(A).$$

Из равенств (*), (**) и (***) следует, что для каждого индекса $i \in I$ верна формула Байеса.

Замечание. Формула Байеса может быть записана также следующим образом:

$$P_A(B_i) = P(B_i) P_{B_i}(A)/P(A),\ P(A) = \sum_{j \in I} P(B_j) P_{B_j}(A).$$

Стандартные задачи, решаемые с помощью формулы Байеса, можно схематически описать следующим образом.

Задача о вероятностях гипотез. *Известны*:

1) вероятности $P(B_i) = \beta_i$ нескольких исключающих друг друга предположений B_i, одно из которых верно;

2) условные вероятности $P_{B_i}(A) = \alpha_i$ события A при условии, что верно предположение B_i.

Какова условная вероятность $P_A(B_i)$ того, что верно предположение B_i при условии, что реализуется событие A?

Решение. Сравнивая условия задачи о вероятностях гипотез и задачи о полной вероятности, видим, что для задачи о вероятностях гипотез можно использовать ту же модель, что и для задачи о полной вероятности. Решение задачи о вероятностях гипотез сводится к применению формулы Байеса для этой модели:

$$P_A(B_i) = P(B_i) P_{B_i}(A)/P(A) = \beta_i \alpha_i/\alpha,$$

$$P(A) = \sum_{j \in I} P(B_j) P_{B_j}(A) = \sum_{j \in I} \beta_j \alpha_j = \alpha$$

для каждого индекса $i \in I$.

Примеры. Рассмотрим несколько примеров, поясняющих использование формулы Байеса.

Задача 1. *Команда составлена из двух отличных, трех хороших и пяти средних стрелков. Каждый отличный стрелок попадает в мишень с вероятностью 0,99, хороший — 0,90 и средний — 0,75. Наугад выбранный из команды стрелок попадает в мишень. Какова вероятность того, что это: 1) отличный; 2) хороший, 3) средний стрелок?*

Решение. Из сказанного при решении задачи 1 пункта 5 ясно, что данная задача сводится к задаче о вероятностях гипотез с вероятностями $\beta_1 = 0{,}2$; $\beta_2 = 0{,}3$; $\beta_3 = 0{,}5$ гипотез (выбирается отличный, хороший, средний стрелок) и условными вероятностями $\alpha_1 = 0{,}99$; $\alpha_2 = 0{,}90$; $\alpha_3 = 0{,}75$ события по этим гипотезам (попадает отличный, хороший, средний стрелок). Используя полученный при решении задачи 1 пункта 5 результат и формулу Байеса, находим, что искомые вероятности соответственно равны:

$$1)\quad P_A(B_1) = \frac{\beta_1 \alpha_1}{\alpha} \approx \frac{0{,}2 \cdot 0{,}99}{0{,}82} \approx 0{,}22;$$

$$2)\quad P_A(B_2) = \frac{\beta_2 \alpha_2}{\alpha} \approx \frac{0{,}3 \cdot 0{,}90}{0{,}82} \approx 0{,}33;$$

$$3)\quad P_A(B_3) = \frac{\beta_3 \alpha_3}{\alpha} \approx \frac{0{,}5 \cdot 0{,}75}{0{,}82} \approx 0{,}45.$$

Замечание. Наиболее вероятно, что наугад выбранный и попавший в мишень стрелок оказывается средним стрелком. Это объясняется тем, что средние стрелки составляют большинство в команде, а разница в квалификации недостаточно велика для того, чтобы компенсировать разницу в числе стрелков.

В то же время вероятность того, что наугад выбранный из команды и попавший в мишень стрелок оказывается средним стрелком, меньше 0,5. Поэтому более надежно предположить, что наугад выбранный из команды и попавший в мишень стрелок является средним либо хорошим стрелком:

$$P_A(B_3+B_2)=P_A(B_3)+P_A(B_2)\approx 0{,}78.$$

Упражнение. Решить задачу 1 для команды из трех отличных, трех хороших и четырех средних стрелков.

Задача 2. *Имеются две книги: русская и английская. Из них наугад выбирается одна книга. Наугад выбранная из нее буква оказывается буквой «е». Какова вероятность того, что эта книга: 1) русская, 2) английская?*

Решение. Из сказанного при решении задачи 2 пункта 5 ясно, что данная задача сводится к задаче о вероятностях гипотез с $\beta_1=0{,}5$ и $\beta_2=0{,}5$ (выбирается русская, английская книга) и $\alpha_1=0{,}072$, $\alpha_2=0{,}131$ (выбранная из русской, английской книги буква оказывается буквой «е»). Используя вычисления, проделанные при решении задачи 2 пункта 5, и формулу Байеса, находим, что искомые вероятности соответственно равны:

$$1)\quad P_A(B_1)=\frac{\beta_1\alpha_1}{\beta_1\alpha_1+\beta_2\alpha_2}=\frac{72}{203}\approx 0{,}35;$$

$$2)\quad P_A(B_2)=\frac{\beta_2\alpha_2}{\beta_1\alpha_1+\beta_2\alpha_2}=\frac{131}{203}\approx 0{,}65.$$

Замечание. Более вероятно, что выбранная книга оказывается английской. Это объясняется тем, что в английском тексте буква «е» встречается чаще, чем в русском.

Упражнение. Решить задачу 2 для двух русских и одной английской книг.

Задача 3. *Множество натуральных чисел {1, ..., 100} разбито на десятки {1, ..., 10}, ..., {91, ..., 100}. Из них наугад выбирается один десяток. Наугад выбранное из него число оказывается простым. Какова вероятность того, что в этом десятке ровно 1, 2, 3, 4 простых числа?*

Решение. Из сказанного при решении задачи 3 пункта 5 ясно, что данная задача сводится к задаче о вероятностях гипотез с $\beta_1=0{,}1$; $\beta_2=0{,}5$; $\beta_3=0{,}2$; $\beta_4=0{,}2$ (выбирается десяток с 1, 2, 3, 4 простыми числами) и $\alpha_1=0{,}1$, $\alpha_2=0{,}2$, $\alpha_3=0{,}3$, $\alpha_4=0{,}4$ (выбранное из десятка с 1, 2, 3, 4 простыми числами число оказывается простым). Используя полученный при решении задачи 3 пункта 5

результат и формулу Байеса, находим, что искомые вероятности соответственно равны:

$$1)\quad P_A(B_1)=\frac{\beta_1\alpha_1}{\alpha}=\frac{0,1\cdot 0,1}{0,25}=0,04;$$

$$2)\quad P_A(B_2)=\frac{\beta_2\alpha_2}{\alpha}=\frac{0,5\cdot 0,2}{0,25}=0,40;$$

$$3)\quad P_A(B_3)=\frac{\beta_3\alpha_3}{\alpha}=\frac{0,2\cdot 0,3}{0,25}=0,24;$$

$$4)\quad P_A(B_4)=\frac{\beta_4\alpha_4}{\alpha}=\frac{0,2\cdot 0,4}{0,25}=0,32.$$

Замечание. Наиболее вероятно, что в выбранном десятке 2 простых числа. Это объясняется тем, что таких десятков больше, чем других. Несколько менее вероятно, что в выбранном десятке 4 простых числа. Хотя таких десятков только два, но в каждом из них наибольшее количество простых чисел.

Более надежно предположить, что в десятке 2 либо 4 простых числа:

$$P_A(B_2+B_4)=P_A(B_2)+P_A(B_4)=0,72.$$

Наименее вероятно, что в выбранном десятке 1 простое число. Такой десяток всего один, и в нем наименьшее количество простых чисел.

Упражнение. Решить задачу 3 для множества натуральных чисел $\{101, \ldots, 200\}$.

§ 5. НЕЗАВИСИМОСТЬ И ЗАВИСИМОСТЬ

При исследовании явлений, определяемых возможными результатами опыта, часто бывает важно установить и измерить зависимость между некоторыми из этих явлений. Поэтому для рассматриваемой вероятностной модели определяется зависимость между событиями, вводятся меры для нее.

1.5. Определение. Независимость и зависимость

Рассмотрим конечную вероятностную модель $(U, \mathscr{A}; p, P)$ и произвольные события A, B для нее.

Определение 7. *События A и B называются независимыми, если верно равенство*

$$P(AB)=P(A)P(B).$$

Если это равенство не верно:

$$P(AB)\neq P(A)P(B),$$

то события A и B называются зависимыми.

Понятия зависимости и независимости поясняют примеры пунктов 1, 2 и 3 § 5 главы 1.

Замечание. Если одно из событий A и B невозможно, то эти события независимы. Например, если $P(B)=0$, то по правилу неравенства $P(AB)=0$ и, следовательно, $P(AB)=P(A)P(B)=0$. Условимся каждое событие, вероятность которого равна 1, называть достоверным. Если одно из событий A и B достоверно, то эти события независимы. Например, если $P(B)=1$, то $P(B')=0$, $P(AB')=0$ и, следовательно,

$$P(AB)=P(AB)+P(AB')=P(A)=P(A)P(B).$$

События A и $B=A$ независимы, если и только если событие A невозможно или достоверно: равенства $P(AA)=P(A)=P(A)P(A)$ верны, если и только если $P(A)=0$ или $P(A)=1$.

Зависимость возможных событий можно описать с помощью условной вероятности.

Предложение. *Если событие B возможно, то события A и B независимы тогда и только тогда, когда верно равенство*

$$P_B(A)=P(A).$$

Доказательство. Если событие B возможно и события A и B независимы, то, используя правило деления, получаем:

$$P_B(A)=P(AB)/P(B)=P(A)P(B)/P(B)=P(A).$$

Обратно, если верно условие высказанного предложения, то, используя правило умножения, имеем:

$$P(AB)=P(B)P_B(A)=P(B)P(A).$$

Предложение доказано.

Следовательно, если событие B возможно, то события A и B зависимы тогда и только тогда, когда

$$P_B(A)\neq P(A).$$

Пример 1. В задаче 1 пункта 4 § 4 события A и B, описывающие соответственно получение числа очков строго большего пяти и попадание в мишень, зависимы:

$$P_B(A)=90/99\neq 90/100=P(A).$$

Пример 2. В задаче 2 пункта 4 § 4 события A и B, описывающие соответственно выбор буквы «а» и выбор гласной буквы, зависимы:

$$P_B(A)=0{,}062/0{,}350\neq 0{,}62=P(A).$$

Пример 3. В задаче 3 пункта 4 § 4 события A и B, описывающие соответственно выбор десятка с нечетным количеством простых чисел и выбор четного десятка, зависимы:

$$P_B(A)=2/5\neq 3/10=P(A).$$

Независимые события A и B рассматриваются в примерах 1 и 2 § 4 главы 2.

Сравнивая условные вероятности $P_{B_i}(A)$ и вероятность $P(A)$ в задачах 1—3 пункта 5 из § 4, посвященного формуле полной вероятности, убеждаемся в зависимости событий B_i и A для каждого индекса i.

2.5. Коэффициенты регрессии

Предлагаемый в качестве меры зависимости коэффициент регрессии учитывает различный характер зависимости одного события от другого.

Рассмотрим конечную вероятностную модель $(U, \mathcal{A}; p, P)$, возможное событие B, дополнительное событие B' для которого тоже возможно, и произвольное событие A для этой модели:

$$P(B),\ P(B') = 1 - P(B) > 0.$$

Определение 8. *Коэффициентом регрессии события A относительно события B называется число*

$$R(A, B) = \frac{P(AB) - P(A)P(B)}{P(B)(1 - P(B))}.$$

Понятие коэффициента регрессии поясняют примеры 1—3 пунктов 1 и 3 § 5 главы 1.

Замечание. Коэффициент регрессии $R(A, B)$ несимметричен. Он измеряет зависимость события A от события B и может быть не равен коэффициенту регрессии

$$R(B, A) = \frac{P(AB) - P(A)P(B)}{P(A)(1 - P(A))},$$

измеряющему зависимость события B от возможного события A, дополнительное событие A' для которого тоже возможно.

Пример. В задаче 2 пункта 4 § 4 событие A описывает появление буквы «а» при выборе наугад буквы из русской книги, а событие B — гласной. Используя проведенные при решении этой задачи вычисления, получаем:

$$R(A, B) = \frac{0{,}062 - 0{,}062 \cdot 0{,}350}{0{,}350\,(1 - 0{,}350)} = \frac{0{,}062}{0{,}350} \approx 0{,}2;$$

$$R(B, A) = \frac{0{,}062 - 0{,}062 \cdot 0{,}350}{0{,}062\,(1 - 0{,}062)} = \frac{0{,}650}{0{,}938} \approx 0{,}7.$$

Рассмотрим некоторые свойства коэффициента регрессии $R(A, B)$ события A относительно события B.

1-е свойство: *события A и B независимы, если и только если $R(A, B) = 0$.*

Для того, чтобы убедиться в этом, достаточно сравнить определения независимости и коэффициента регрессии, заметив, что равенства

$$P(AB) = P(A)P(B) \quad \text{и} \quad P(AB) - P(A)P(B) = 0$$

эквивалентны.

2-е свойство: $R(A, B) = P_B(A) - P_{B'}(A)$.

Как и в пункте 3 § 5 главы 1, используя правила деления и вычитания, получаем:

$$(*)\quad P_B(A) - P_{B'}(A) = \frac{P(AB)}{P(B)} - \frac{P(A) - P(AB)}{1 - P(B)} = R(A, B).$$

3-е свойство: $|R(A, B)| \leqslant 1$.
Так как

$$(**)\quad 0 \leqslant P_B(A) \leqslant 1, \quad 0 \leqslant P_{B'}(A) \leqslant 1,$$

то, вычитая вторые неравенства из первых, получаем:

$$-1 \leqslant P_B(A) - P_{B'}(A) \leqslant 1.$$

Отсюда и из равенств (*) следует доказываемое неравенство.

4-е свойство: $R(AB) = 1$, *если и только если*

$$P(A) = P(B) = P(AB).$$

Из равенств (*) и неравенств (**) следует, что равенство

$$R(A, B) = 1$$

эквивалентно равенствам

$$P_B(A) = 1, \quad P_{B'}(A) = 0.$$

Из правила умножения вытекает, что эти равенства эквивалентны соответственно равенствам

$$P(AB) = P(B),\ P(AB') = 0.$$

Из правила сложения следует, что второе из этих равенств эквивалентно равенству

$$P(A) = P(AB) \quad (P(A) = P(AB) + P(AB')).$$

Поэтому равенства $P_B(A) = 1$, $P_{B'}(A) = 0$ эквивалентны равенствам

$$P(A) = P(B) = P(AB).$$

Замечание. Используя равенства $P_B(A) = 1$ и $P_{B'}(A) = 0$, можно сформулировать 4-е свойство коэффициента регрессии следующим образом: $R(A, B) = 1$, если и только если событие A достоверно при условии B и невозможно при условии B'.

5-е свойство: $R(A, B') = -R(A, B)$.

В самом деле, используя правила вычитания и дополнения, получаем:

$$R(A, B') = \frac{P(AB') - P(A)P(B')}{P(B')(1 - P(B'))} = \frac{P(A) - P(AB) - P(A) + P(A)P(B)}{(1 - P(B))P(B)} =$$
$$= -R(A, B).$$

Примеры. $R(B, B) = 1$, $R(B, B') = -1$.

3.5. Коэффициент корреляции

Коэффициент корреляции является симметричной мерой зависимости событий.

Рассмотрим конечную вероятностную модель (U, $\mathcal{A}$; p, P) и возможные события A и B, дополнительные события A' и B' для которых тоже возможны:

$$P(A),\ P(A')=1-P(A),\ P(B),\ P(B')=1-P(B)>0.$$

Определение 9. *Коэффициентом корреляции событий A и B называется число*

$$K(A,B)=\frac{P(AB)-P(A)P(B)}{\sqrt{P(A)(1-P(A))P(B)(1-P(B))}}.$$

Понятие коэффициента корреляции поясняют примеры 1—3 пунктов 1 и 4 § 5 главы 1.

Симметричность коэффициента корреляции следует из его определения:

$$K(A,\ B)=K(B,\ A).$$

Рассмотрим некоторые свойства коэффициента корреляции $K(A,\ B)$ событий A и B.

1-е свойство: *события A и B независимы, если и только если $K(A,\ B)=0$.*

Для того, чтобы убедиться в этом, достаточно сравнить определения независимости и коэффициента корреляции, заметив, что равенства

$$P(AB)=P(A)P(B) \quad \text{и} \quad P(AB)-P(A)P(B)=0$$

эквивалентны.

Условимся называть средним геометрическим чисел a и b число $+\sqrt{ab}$, если a и b положительны, и число $-\sqrt{ab}$, если a и b отрицательны.

2-е свойство: *коэффициент корреляции $K(A,\ B)$ событий A и B равен среднему геометрическому их коэффициентов регрессии $R(A,\ B)$ и $R(B,\ A)$.*

Сравнивая определения коэффициента корреляции $K(A,\ B)$ и коэффициентов регрессии $R(A,\ B)$, $R(B,\ A)$, убеждаемся в том, что верно равенство

$$K(A,\ B)=\pm\sqrt{R(A,\ B)R(B,\ A)},$$

в котором знак выбирается равным общему знаку коэффициентов регрессии.

Пример. В задаче 2 пункта 4 § 4 событие A описывает появление буквы «а» при выборе наугад буквы из русской книги, а событие B — некоторой гласной буквы. Используя проведенные в пункте 2 вычисления, получаем:

$$K(A,\ B)=\pm\sqrt{(62/350)\ (650/938)}\approx 0{,}4.$$

3-е свойство: $|K(A, B)| \leqslant 1$.

4-е свойство: $K(A, B) = 1$, *если и только если*

$$P(A) = P(B) = P(AB).$$

5-е свойство: $K(A, B') = -K(A, B)$.

Все эти свойства следуют из 2-го свойства коэффициента корреляции и соответствующих свойств коэффициентов регрессии.

Примеры. $K(A, A) = 1$, $K(A, A') = -1$.

Замечание. Примеры с вычислением коэффициентов регрессии и коэффициентов корреляции для появления буквы «а» (событие A) и появления некоторой гласной буквы (событие B) в русском тексте показывают, что при истолковании значений этих коэффициентов нужно соблюдать осторожность. Реализация события A автоматически означает реализацию события B. В то же время $R(B, A) \approx 0{,}7$ и $K(A, B) \approx 0{,}4$.

4.5. Коэффициент связи

Коэффициент связи, как и коэффициент корреляции, является симметричной мерой зависимости событий.

Рассмотрим конечную вероятностную модель $(U, \mathscr{A}; p, P)$ и события A и B такие, что

$$P(AB)P(A'B') + P(AB')P(A'B) > 0.$$

Определение 10. *Коэффициентом связи событий A и B называется число*

$$Q(A, B) = \frac{P(AB) - P(A)P(B)}{P(AB)P(A'B') + P(AB')P(A'B)}.$$

Симметричность коэффициента связи следует из определения:

$$Q(A, B) = Q(B, A).$$

Рассмотрим некоторые свойства коэффициента связи $Q(A, B)$ событий A и B.

1-е свойство: *события A и B независимы, если и только если $Q(A, B) = 0$.*

Это свойство проверяется так же, как аналогичные свойства коэффициентов корреляции и регрессии.

2-е свойство: $$Q(A, B) = \frac{P(AB)P(A'B') - P(AB')P(A'B)}{P(AB)P(A'B') + P(AB')P(A'B)}.$$

Действительно, используя правила вычитания и дополнения, получаем:

$$P(AB)P(A'B') = P(AB)P(B' - AB') =$$
$$= P(AB)P(B') - P(AB)P(AB') =$$
$$= P(AB) - P(AB)P(B) - P(AB)P(AB'),$$

$$P(AB')P(A'B) = P(AB')P(B-AB) =$$
$$= P(AB')P(B) - P(AB')P(AB) =$$
$$= P(A)P(B) - P(AB)P(B) - P(AB)P(AB'),$$

откуда

$$P(AB)P(A'B') - P(AB')P(A'B) = P(AB) - P(A)P(B).$$

3-е свойство: $|Q(A, B)| \leqslant 1$.

Следует из 2-го свойства:

$$|P(AB)P(A'B') - P(AB')P(A'B)| \leqslant$$
$$\leqslant P(AB)P(A'B') + P(AB')P(A'B).$$

4-е свойство: $Q(A, B) = 1$, *если и только если* $P(A) = P(AB)$ *или* $P(B) = P(AB)$.

Из 2-го свойства следует, что равенство

$$Q(A, B) = 1$$

эквивалентно равенству

$$P(AB')P(A'B) = 0.$$

Это равенство верно, если и только если верно хотя бы одно из равенств

$$P(AB') = 0, \; P(A'B) = 0.$$

Из правила разности вытекает, что эти равенства эквивалентны соответственно равенствам

$$P(A) - P(AB) = 0, \; P(B) - P(AB) = 0,$$

т. е. равенствам

$$P(A) = P(AB), \; P(B) = P(AB).$$

Примеры. Если $A \subseteq B$ или $B \subseteq A$, то $Q(A, B) = 1$.

5-е свойство: $Q(A, B') = -Q(A, B)$.

Следует из 2-го свойства.

Пример. В задаче 2 пункта 4 § 4 событие A описывает появление буквы «а» при выборе наугад буквы из русской книги, а событие B — некоторой гласной буквы. Так как $A \subseteq B$, то, по 4-у свойству,

$$Q(A, B) = 1.$$

Коэффициент связи по абсолютной величине всегда больше коэффициентов регрессии и корреляции.

Лемма. $|R(A, B)| \leqslant |Q(A, B)|$.

Доказательство. Предполагается, что события A и B удовлетворяют условиям, при которых определялись коэффициент корреляции и коэффициент связи. Разделив числитель и знаменатель $Q(A, B)$ на $P(B)P(B')$ и использовав правило деления для вероятности, получаем:

$$Q(A, B) = \frac{R(A, B)}{P_B(A)P_{B'}(A') + P_B(A')P_{B'}(A)}.$$

По правилам дополнения и сложения

$$P_{B'}(A') = 1 - P_{B'}(A), \quad P_{B'}(A) = 1 - P_{B'}(A'),$$
$$P_B(A) + P_B(A') = 1.$$

Следовательно,

$$0 \leqslant P_B(A) P_{B'}(A') + P_B(A') P_{B'}(A) = P_B(A) - P_B(A) P_{B'}(A) +$$
$$+ P_B(A') - P_B(A') P_{B'}(A') = 1 - (P_B(A) P_{B'}(A) + P_B(A') P_{B'}(A')) \leqslant 1.$$

Поэтому

$$Q(A, B) = \frac{R(A, B)}{1 - (P_B(A) P_{B'}(A) + P_B(A') P_{B'}(A'))},$$

причем знаменатель в правой части положителен и меньше единицы. Отсюда вытекает доказываемое неравенство.

Предложение. $|K(A, B)| \leqslant |Q(A, B)|$.

Доказательство. Так как коэффициент связи симметричен, то из леммы вытекает, что

$$|R(B, A)| \leqslant |Q(B, A)| = |Q(A, B)|,$$

поэтому

$$|K(A, B)| = +\sqrt{R(A, B) R(B, A)} \leqslant \sqrt{|Q(A, B)|^2} = |Q(A, B)|,$$

что и требовалось доказать.

5.5. Независимость семейства событий

Рассмотрим конечную вероятностную модель $(U, \mathscr{A}; p, P)$ и семейство $(C_i)_{i \in I}$ событий C_i, имеющее конечное множество индексов I.

Определение 11. *События C_i семейства $(C_i)_{i \in I}$ называются независимыми, если для каждого множества $K \subseteq I$ верно равенство*

$$P\left(\bigcap_{i \in K} C_i\right) = \prod_{i \in K} P(C_i).$$

Если для некоторого множества $L \subseteq I$ это равенство не верно, то события C_i семейства $(C_i)_{i \in I}$ называются зависимыми.

Примеры 1, 2 и лемма пункта 1 § 5 главы 2 поясняют понятие независимости семейства событий.

Рассмотрим семейство $(\mathscr{C}_i)_{i \in I}$ классов $\mathscr{C}_i$ событий, имеющее конечное множество индексов I. Выбирая по одному событию C_i из каждого класса $\mathscr{C}_i$, можно составить семейство $(C_i)_{i \in I}$ событий, имеющее то же множество индексов I. События C_i могут быть независимы и могут быть зависимы.

Определение 12. *Классы $\mathscr{C}_i$ семейства $(\mathscr{C}_i)_{i \in I}$ называются независимыми, если произвольно выбранные по одному из каждого класса $\mathscr{C}_i$ события C_i независимы.*

Если существует семейство $(C_i)_{i\in I}$, события $C_i\in\mathscr{C}_i$ которого зависимы, то и классы $\mathscr{C}_i$ называются зависимыми.

Примеры 1, 2 и теорема пункта 2 § 5 главы 2 поясняют понятие независимости классов событий.

Рассмотрим семейство $(C_i)_{i\in I}$ событий C_i, имеющее конечное множество индексов I.

Определение 11′. *События C_i семейства $(C_i)_{i\in I}$ называются попарно независимыми, если и только если для каждых двух различных индексов i и j события C_i и C_j независимы*:

$$P(C_iC_j)=P(C_i)P(C_j).$$

Сравнивая определения 11 и 11′, видим, что если события C_i семейства $(C_i)_{i\in I}$ независимы, то они попарно независимы. Обратное предложение не верно. Это показывает следующий

Пример Бернштейна. Правильный тетраэдр раскрашен следующим образом: одна грань в красный цвет, вторая — в синий, третья — в зеленый, четвертая — в эти три цвета. Наугад выбирается грань тетраэдра и отмечается ее окраска.

Такой опыт описывается стандартной моделью Лапласа $n=4$ равновозможных исходов. Исходы 1, 2, 3 описывают соответственно выбор красной, синей, зеленой граней, а исход 4 — трехцветной.

События $C_1=\{1, 4\}$, $C_2=\{2, 4\}$, $C_3=\{3, 4\}$ описывают появление соответственно красного, синего, зеленого цвета, события $C_1C_2=C_1C_3=C_2C_3=\{4\}$ — появление двух цветов, событие $C_1C_2C_3=\{4\}$ — появление трех цветов. (В то же время все эти пересечения описывают выбор трехцветной грани).

Рассмотрим семейство $(C_i)_{i\in I}$ событий C_i, имеющее множество индексов $I=\{1, 2, 3\}$. События C_i попарно независимы:

$$P(C_1C_2)=P(C_1C_3)=P(C_2C_3)=1/4;$$

$$P(C_1)P(C_2)=P(C_1)P(C_3)=P(C_2)P(C_3)=$$

$$=1/2\cdot 1/2=1/4.$$

В то же время события C_i зависимы:

$$P(C_1C_2C_3)=1/4\neq 1/2\cdot 1/2\cdot 1/2=P(C_1)P(C_2)P(C_3).$$

Задачи по конечным вероятностным моделям собраны в главе 2 части III.

Часть II

СЛУЧАЙНЫЕ ПЕРЕМЕННЫЕ

Часто задача сводится к исследованию некоторой функции на множестве исходов. Такие функции называются случайными переменными. Основной характеристикой случайной переменной является взвешенное с помощью элементарных вероятностей среднее ее значений. Нередко задача сводится к вычислению такого среднего. Для решения этой задачи используются различные правила вычисления средних. С помощью среднего определяются другие важные характеристики случайных переменных: дисперсия, информация, энтропия.

Глава 1

РАСПРЕДЕЛЕНИЕ

Вероятностные свойства случайной переменной определяются ее распределением, т. е. вероятностью принимать значения из данного множества чисел.

При первом чтении рекомендуется ограничиться § 1, 2 и началом § 3.

§ 1. ПРИМЕРЫ

Рассмотрим несколько примеров, разъясняющих понятия *случайной переменной* и ее *распределения*. Эти примеры подготавливают соответствующие формальные определения.

1.1. Пример 1. Игра с костью

Игральная кость подбрасывается один раз. Если выпадает четное число очков, то игрок выигрывает 1 рубль, а если нечетное — проигрывает. Какова вероятность того, что игрок выигрывает 1 рубль?

Как было подробно объяснено в примере 1 из § 1 главы 1 части I, подбрасывание кости один раз описывается стандартной моделью Лапласа $n=6$ равновероятных исходов 1, 2, 3, 4, 5, 6.

1.1.1. Случайная переменная

Выигрыш игрока удобно описывать функцией f на множестве исходов $U=\{1, 2, 3, 4, 5, 6\}$, значение $f(u)$ которой для каждого исхода u определяется таблицей

u	1	2	3	4	5	6
$f(u)$	-1	$+1$	-1	$+1$	-1	$+1$

Игрок выигрывает $f(u)$ рублей, если реализуется результат, описываемый исходом u (выпадает u очков). Реализация этого результата зависит от случая. Поэтому функцию f называют *случайной переменной.* Случайная переменная f принимает значение $f(u)=+1$ при $u=2, 4, 6$ и значение $f(u)=-1$ при $u=1, 3, 5$. В рассматриваемой модели $(U, p; \mathscr{A}, P)$ задача сводится к вычислению вероятности $P(A)$ события $A=f^{-1}(+1)=\{u:f(u)=+1\}=\{2, 4, 6\}$.

2.1.1. Новая модель

Можно использовать для решения задачи другую вероятностную модель. Так как нас интересует значение случайной переменной f, то в качестве нового *множества исходов* естественно взять множество значений

$$X=f[U]=\{+1, -1\}$$

случайной переменной f.

3.1.1. Элементарное распределение

Новый исход $x=+1$ и прежнее событие $f^{-1}(X)=f^{-1}(+1)=\{u:f(u)=+1\}=\{2, 4, 6\}$ описывают одно и то же явление — выпадение *четного* числа очков. Поэтому исход $x=+1$ и событие $f^{-1}(x)=f^{-1}(+1)$ должны иметь одну и ту же *меру реализуемости.* Исходу $x=+1$ приписывается элементарная вероятность $q(x)=q(+1)$, равная вероятности $P(f^{-1}(x))=P(f^{-1}(+1))$ события $f^{-1}(x)=f^{-1}(+1)$:

$$q(+1)=P(f^{-1}(+1))=P(\{u:f(u)=+1\})=P(\{2, 4, 6\})=1/2.$$

Аналогично исходу $x=-1$ приписывается элементарная вероятность $q(x)=q(-1)$, равная вероятности $P(f^{-1}(x))=P(f^{-1}(-1))$ события $f^{-1}(x)=f^{-1}(-1)$:

$$q(-1)=P(f^{-1}(-1))=P(\{u:f(u)=-1\})=P(\{1, 3, 5\})=1/2.$$

Так оценивается реализуемость *нечетного* числа очков.

Числа $q(+1)$ и $q(-1)$ *положительны* и в сумме равны *единице*. Функция q на множестве исходов X, значение $q(x)$ которой для каждого исхода x из X определяются таблицей

x	$+1$	-1
$q(x)$	$\frac{1}{2}$	$\frac{1}{2}$

является *элементарной вероятностью* для новой модели. Элементарная вероятность q называется *элементарным распределением* для случайной переменной f.

4.1.1. Распределение

Множество исходов X и элементарная вероятность q определяют вероятностную модель $(X, q; \mathscr{B}, Q)$. В этой модели *алгебра событий* образована классом

$$\mathscr{B}=\{O, \{+1\}, \{-1\}, X\}$$

всех частей множества X с определенными для них объединением, пересечением и дополнением. А *вероятность* Q на классе событий $\mathscr{B}$ имеет значение

$$Q(B) = \sum_{x\in B} q(x)$$

для каждого события B из $\mathscr{B}$:

$$Q(O)=0,\ Q(\{+1\})=1/2,\ Q(\{-1\})=1/2,\ Q(X)=1.$$

Вероятность Q называется *распределением* для случайной переменной f.

Можно определить распределение Q, используя вероятность P. Для каждого нового события B рассмотрим его прообраз

$$f^{-1}(B)=\{u: f(u)\in B\}:$$

$$f^{1-}(O)=O,\ f^{-1}(\{+1\})=\{2, 4, 6\},\ f^{-1}(\{-1\})=\{1, 3, 5\},\ f^{-1}(X)=U.$$

Из определений распределения Q и элементарного распределения q для случайной переменной f следует, что

$$Q(B)=P(f^{-1}(B)).$$

Новое событие B и прежнее событие $A=f^{-1}(B)$ описывают одно и то же явление. Поэтому они имеют одну и ту же вероятность.

Таким образом, вероятностная модель $(U, p; \mathscr{A}, P)$ и случайная переменная f на множестве U определяют *новую вероятностную модель* $(X, q; \mathscr{B}, Q)$. В этой новой модели задача сводится к вычислению вероятности $Q(B)$ события $B=\{+1\}$:

$$Q(B)=Q(\{+1\})=q(+1)=P(f^{-1}(+1))=P(A)=1/2.$$

Заметим, что для решения задачи достаточно знать только распределение Q случайной переменной f, а не саму эту переменную.

2.1. Пример 2. Игра с монетой

Симметричная монета подбрасывается два раза. Если при первом подбрасывании выпадает герб, то игрок выигрывает 5 копеек, а если цифра — проигрывает. Так же при втором подбрасывании. Какова вероятность того, что игрок не проигрывает?

Как было подробно объяснено в примере 3 из § 1 главы 1 части I, подбрасывание симметричной монеты два раза описывается моделью Бернулли $n=2$ испытаний с вероятностью успеха $a=1/2$.

1.2.1. Случайная переменная

Выигрыш игрока при первом подбрасывании удобно описывать функцией f_1 на множестве исходов $U=\{11, 00, 01, 10\}$, значение $f_1(u)$ которой для каждого исхода u определяется таблицей

u	11	10	01	00
$f_1(u)$	+5	+5	—5	—5

.

Выигрыш игрока при втором подбрасывании удобно описывать функцией f_2 на множестве U, значение $f_2(u)$ которой для каждого исхода u определяется таблицей

u	11	10	01	00
$f_2(u)$	+5	—5	+5	—5

.

Наконец, общий выигрыш игрока удобно описывать функцией $f=f_1+f_2$ на множестве U, значение $f(u)=f_1(u)+f_2(u)$ которой для каждого исхода u определяется таблицей

u	11	10	01	00
$f(u)$	+10	0	0	—10

.

Игрок выигрывает $f(u)$ рублей, если реализуется результат, описываемый исходом u. Реализация этого результата зависит от случая. Поэтому функцию f называют *случайной переменной.* Функции f_1 и f_2 тоже называют случайными переменными.

Случайная переменная f принимает значение $f(u)=+10$ при $u=11$, значение $f(u)=0$ при $u=10$, 01 и значение $f(u)=-10$ при $u=00$. В рассматриваемой модели $(U, p; \mathcal{A}, P)$ задача сводится к вычислению вероятности $P(A)$ события $A=f^{-1}(\{+10, 0\}=$ $=\{u:f(u)\in\{+10, 0\}\}=\{11, 10, 01\}$:

$$P(A)=P(11)+P(10)+P(01)=3/4.$$

2.2.1. Новая модель

Можно использовать для решения задачи другую вероятностную модель. Так как нас интересуют значения случайной переменной f, то в качестве нового *множества исходов* естественно взять множество значений

$$X = f(U) = \{+10,\ 0,\ -10\}$$

случайной переменной f.

3.2.1. Элементарное распределение

Новый исход $x = +10$ и прежнее событие $f^{-1}(x) = f^{-1}(+10) = \{u : f(u) = +10\} = \{11\}$ описывают одно и то же явление — выпадение двух гербов. Поэтому в качестве *меры реализуемости* исходу $x = +10$ приписывается элементарная вероятность $q(x) = q(+10)$, равная вероятности $P(f^{-1}(x)) = P(f^{-1}(+10))$ события $f^{-1}(x) = f^{-1}(+10)$:

$$q(+10) = P(f^{-1}(+10)) = P(\{u : f(u) = +10\}) = P(\{11\}) = 1/4.$$

Аналогично,

$$q(0) = P(f^{-1}(0)) = P(\{u : f(u) = 0\}) = P(\{10,\ 01\}) = 1/2,$$

$$q(-10) = P(f^{-1}(-10)) = P(\{u : f(u) = -10\}) = P(\{00\}) = 1/4.$$

Числа $q(+10)$, $q(0)$, $q(-10)$ положительны и в сумме равны единице. Функция q на множестве исходов X, значение $q(x)$ которой для каждого исхода x из X определяется таблицей

x	$+10$	0	-10
$q(x)$	$1/4$	$1/2$	$1/4$

является элементарной вероятностью для новой модели. Элементарная вероятность q называется *элементарным распределением* для случайной переменной f.

4.2.1. Распределение

Множество исходов X и элементарная вероятность q определяют вероятностную модель $(X,\ q;\ \mathscr{B},\ Q)$. В этой модели алгебра событий образована классом

$$\mathscr{B} = \{O,\ \{+10\},\ \{-10\},\ \{0\},\ \{+10,\ 0\},\ \{+10,\ -10\},\ \{0,\ -10\},\ X\}$$

всех частей множества X с определенными для них объединением, пересечением и дополнением. А вероятность Q на классе событий $\mathscr{B}$ имеет значение

$$Q(B) = \sum_{x \in B} q(x)$$

для каждого события B из $\mathscr{B}$:

B	O	$\{+10\}$	$\{0\}$	$\{-10\}$	$\{+10, 0\}$	$\{+10, -10\}$	$\{0, -10\}$	X
$Q(B)$	0	1/4	1/2	1/4	3/4	1/2	3/4	1

Вероятность Q называется *распределением* для случайной переменной f.

Можно определить распределение Q, используя вероятность P. Для каждого нового события B рассмотрим его прообраз $f^{-1}(B)=\{u: f(u)\in B\}$:

B	O	$\{+10\}$	$\{0\}$	$\{+10\}$	$\{+10, 0\}$	$\{+10, -10\}$	$\{0, -10\}$	X
$f^{-1}(B)$	0	$\{11\}$	$\{10, 01\}$	$\{00\}$	$\{11, 10, 01\}$	$\{11, 00\}$	$\{10, 01, 00\}$	U

Из определений распределения Q и элементарного распределения q для случайной переменной f следует, что

$$Q(B)=P(f^{-1}(B)).$$

Новое событие B и прежнее событие $A=f^{-1}(B)$ описывают одно и то же явление. Поэтому они имеют одну и ту же вероятность.

Таким образом, вероятностная модель $(U, p; \mathscr{A}, P)$ и случайная переменная f на множестве U определяют *новую вероятностную модель* $(X, q; \mathscr{B}, Q)$. В этой новой модели задача сводится к вычислению вероятности $Q(B)$ события $B=\{+10, 0\}$:

$$Q(B)=Q(\{+10, 0\})=q(+10)+q(0)=P(f^{-1}(+10))+$$
$$+P(f^{-1}(0))=P(f^{-1}(\{+10, 0\}))=3/4.$$

Как и в примере 1, для решения задачи достаточно знать только распределение Q случайной переменной f, а не саму эту переменную.

Замечание. Элементарные распределения q_1, q_2 и распределения Q_1, Q_2 для случайных переменных f_1, f_2 определяются соответственно таблицами

x_1	$+5$	-5
$q(x_1)$	1/2	1/2

;

x_2	$+5$	-5
$q(x_2)$	1/2	1/2

;

B_1	O	$\{+5\}$	$\{-5\}$	X_1
$Q_1(B_1)$	0	1/2	1/2	1

;

B_2	O	$\{+5\}$	$\{-5\}$	X_2
$Q_2(B_2)$	0	1/2	1/2	1

.

Различные случайные переменные f_1 и f_2 $(f_1(10)\neq f_2(10),$ $f_1(01)\neq f_2(01))$ имеют одинаковые элементарные распределения $q_1=q_2$ и одинаковые распределения $Q_1=Q_2$.

3.1. Пример 3. Выстрел по мишени

Производится один выстрел по стандартной мишени. Известны вероятности получения каждого числа очков. Какова вероятность того, что получается число очков, строго большее пяти?

Как было подробно объяснено в примере 1 из § 1 главы 3 части I, один выстрел по стандартной мишени описывается вероятностной моделью $(U, p; \mathcal{A}, P)$, определяемой множеством исходов

$$U = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

и элементарной вероятностью p на множестве U, которая предполагается известной. Пусть, например, значение $p(u)$ элементарной вероятности p для каждого исхода u определяется таблицей

u	0	1	2	3	4	5	6	7	8	9	10
$p(u)$	1/100	1/100	1/100	1/100	1/100	1/20	1/20	1/20	1/10	1/5	1/2

1.3.1. Случайная переменная

Получаемое число очков удобно описывать тождественной функцией f на множестве исходов U, значение $f(u)$ которой для каждого исхода u равно этому исходу:

$$f(u) = u.$$

Реализация результата, описываемого исходом u (получение u очков), зависит от случая. Поэтому функцию f называют *случайной переменной*. В рассматриваемой модели задача сводится к вычислению вероятности $P(A)$ события $A = \{6, 7, 8, 9, 10\}$:

$$P(A) = p(6) + p(7) + p(8) + p(9) + p(10) = 9/10.$$

2.3.1 Новая модель

Множество значений

$$X = f(U) = U$$

случайной переменной X равно множеству исходов U.

Для каждого исхода $x = u$ верно равенство

$$f^{-1}(x) = \{u\}.$$

Поэтому

$$q(x) = P(f^{-1}(x)) = P(\{u\}) = p(u).$$

Элементарное распределение q для тождественной случайной переменной f равно элементарной вероятности p.

Для каждого события $B = A$ верно равенство

$$f^{-1}(B) = A,$$

поэтому

$$Q(B) = P(f^{-1}(B)) = P(A).$$

Распределение Q для тождественной случайной переменной f равно вероятности P.

Вообще, *новая вероятностная модель* $(X, q; \mathscr{B}, Q)$, определяемая по аналогии с примерами 1 и 2 рассматриваемой вероятностной моделью $(U, p; \mathscr{A}, P)$ и тождественной случайной переменной f, совпадает с прежней:

$$(U, p; \mathscr{A}, P) = (X, q; \mathscr{B}, Q),$$

т. е. $U=X$, $p=q$, $\mathscr{A}=\mathscr{B}$, $P=Q$.

Задача сводится к уже решенной в примере 1 из § 1 главы 3 части I.

Как и в примерах 1, 2, для решения задачи достаточно знать только распределение Q случайной переменной f, а не саму эту переменную.

§ 2. ОПРЕДЕЛЕНИЯ

Случайной переменной называется числовая функция на множестве исходов. Каждая случайная переменная определяет некоторую элементарную вероятность на множестве своих значений. Эта элементарная вероятность называется *элементарным распределением* случайной переменной. Элементарное распределение определяет вероятность на классе всех частей множества значений случайной переменной. Эта вероятность называется *распределением* случайной переменной. Распределение описывает вероятностные свойства случайной переменной. Различные случайные переменные могут иметь одинаковые распределения.

Во всех пунктах этого параграфа рассматривается конечная вероятностная модель $(U, p; \mathscr{A}, P)$.

1.2. Случайная переменная

Определение 1. *Случайной переменной называется каждая числовая функция f, определенная на множестве исходов U.*

Часто случайная переменная f определяется таблицей своих значений: в верхней строке таблицы записываются исходы u, в нижней под каждым исходом u — значение $f(u)$ случайной переменной f для u.

Примеры. В примерах 1—3 соответственно:

u	1	2	3	4	5	6
$f(u)$	−1	+1	−1	+1	−1	+1

;

u	11	10	01	00
$f(u)$	+10	0	0	−10

;

u	0	1	2	3	4	5	6	7	8	9	10
$f(u)$	0	1	2	3	4	5	6	7	8	9	10

.

Для случайных переменных, как и для любых числовых функций, определяются сложение, умножение и умножение на число. Суммой, произведением и произведением на число c случайных переменных f и g на множестве исходов U являются случайные переменные $f+g$, $f\cdot g$ и $c\cdot f$ со значениями соответственно

$$f(u)+g(u),\ f(u)\cdot g(u),\ c\cdot f(u)$$

для каждого исхода u.

Индикаторы. Для каждого исхода u определим случайную переменную i_u, значение которой для каждого исхода v равно 1, если $v=u$, и равно 0, если $v\neq u$:

$$i_u(v)=\begin{cases}1 & (v=u),\\ 0 & (v\neq u).\end{cases}$$

Случайная переменная i_u называется *индикатором исхода* u.

Для каждого события A определим случайную переменную

$$i_A=\sum_{u\in A} i_u,$$

равную сумме индикаторов исходов u, составляющих событие A. Значение $i_A(v)$ случайной переменной i_A для каждого исхода v равно 1, если v принадлежит A, и равно 0, если v не принадлежит A:

$$i_A(v)=\begin{cases}1 & (v\in A),\\ 0 & (v\notin A).\end{cases}$$

Случайная переменная i_A называется *индикатором события* A. Индикаторы исходов и событий используются в следующих примерах.

1.1.2. Пример 1

Рассмотрим стандартную модель Лапласа $n=6$ равновероятных исходов 1, 2, 3, 4, 5, 6, которая описывает подбрасывание игральной кости один раз. Таблицы значений индикаторов i_1, i_2, i_3, i_4, i_5, i_6 имеют вид:

v	1	2	3	4	5	6
$i_1(v)$	1	0	0	0	0	0

;

v	1	2	3	4	5	6
$i_2(v)$	0	1	0	0	0	0

;...;

v	1	2	3	4	5	6
$i_6(v)$	0	0	0	0	0	1

.

Каждая случайная переменная f на множестве $U=\{1, 2, 3, 4, 5, 6\}$ равна некоторой *линейной комбинации индикаторов*. Индикатор i_u исхода u умножается на значение $f(u)$ случайной переменной f для исхода u, и полученные произведения складываются:

$$f=\sum_{1\leqslant u\leqslant 6} i_u f(u).$$

В этой сумме для каждого исхода v значения всех слагаемых, кроме одного, равны 0:

$$f(1) = \sum_{1 \leqslant u \leqslant 6} i_u(1) f(u) =$$
$$= 1 \cdot f(1) + 0 \cdot f(2) + 0 \cdot f(3) + 0 \cdot f(4) + 0 \cdot f(5) + 0 \cdot f(6),$$
$$f(2) = \sum_{1 \leqslant u \leqslant 6} i_u(2) f(u) =$$
$$= 0 \cdot f(1) + 1 \cdot f(2) + 0 \cdot f(3) + 0 \cdot f(4) + 0 \cdot f(5) + 0 \cdot f(6),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$f(6) = \sum_{1 \leqslant u \leqslant 6} i_u(6) f(u) =$$
$$= 0 \cdot f(1) + 0 \cdot f(2) + 0 \cdot f(3) + 0 \cdot f(4) + 0 \cdot f(5) + 1 \cdot f(6).$$

Например, случайная переменная f, рассматривавшаяся в примере 1 из § 1, представляется равенством

$$f = -i_1 + i_2 - i_3 + i_4 - i_5 + i_6.$$

Таблицы значений индикаторов $i_A = i_{\{2,4,6\}}$ и $i_{A'} = i_{\{1,3,5\}}$ событий $A = \{2, 4, 6\}$ и $A' = \{1, 3, 5\}$ имеют вид:

v	1	2	3	4	5	6
$i_A(v)$	0	1	0	1	0	1

;

v	1	2	3	4	5	6
$i_{A'}(v)$	1	0	1	0	1	0

.

Собрав индикаторы с одинаковыми коэффициентами, можно получить следующее представление для случайной переменной f в примере 1 из § 1:

$$f = i_{\{2,4,6\}} - i_{\{1,3,5\}}.$$

2.1.2. Пример 2

Рассмотрим модель Бернулли $n = 2$ испытаний с вероятностью успеха $a = 1/2$, которая описывает подбрасывание монеты два раза. Таблицы значений индикаторов i_{11}, i_{10}, i_{01}, i_{00} имеют вид:

v	11	10	01	00
$i_{11}(v)$	1	0	0	0

;

v	11	10	01	00
$i_{10}(v)$	0	1	0	0

;

v	11	10	01	00
$i_{01}(v)$	0	0	1	0

;

v	11	10	01	00
$i_{00}(v)$	0	0	0	1

.

Случайные переменные f_1, f_2 и f в примере 2 из § 1 представляются следующими линейными комбинациями индикаторов исходов:

$$f_1 = 5 \cdot i_{11} + 5 \cdot i_{10} - 5 \cdot i_{01} - 5 \cdot i_{00};$$

$$f_2 = 5 \cdot i_{11} - 5 \cdot i_{10} + 5 \cdot i_{01} - 5 \cdot i_{00}; \quad f = 10 \cdot i_{11} - 10 \cdot i_{00}.$$

Таблицы значений для индикаторов $i_{У_1}, i_{H_1}, i_{У_2}, i_{H_2}$ событий $У_1 = \{11, 10\}, H_1 = \{01, 00\}, У_2 = \{11, 01\}, H_2 = \{10, 00\}$ имеют вид:

v	11	10	01	00
$i_{У_1}(v)$	1	1	0	0

;

v	11	10	01	00
$i_{H_1}(v)$	0	0	1	1

;

v	11	10	01	00
$i_{У_2}(v)$	1	0	1	0

;

v	11	10	01	00
$i_{H_2}(v)$	0	1	0	1

.

Заметим, что $i_{H_1} \cdot i_{H_2} = i_{H_1 H_2} = i_{00}$, $i_{У_1} \cdot i_{У_1} = i_{У_1 У_1} = i_{11}$, $i_{У_1} \cdot i_{H_2} = i_{У_1 H_2} = i_{10}$, $i_{H_1} \cdot i_{У_2} = i_{H_1 У_2} = i_{01}$. Кроме того, если обозначить символом 1 постоянную со значением единица,

$$i_{H_1} = 1 - i_{У_1}, \quad i_{H_2} = 1 - i_{У_2}.$$

Случайные переменные f_1, f_2 и f в примере 2 из § 1 представляются следующими линейными комбинациями индикаторов событий:

$$f_1 = 5 \cdot i_{У_1} - 5 \cdot i_{H_1}, \quad f_2 = 5 \cdot i_{У_2} - 5 \cdot i_{H_2}, \quad f = 10 \cdot i_{У_1 У_2} - 10 \cdot i_{H_1 H_2}.$$

3.1.2. Пример 3

В примере 3 из § 1 получаемое при выстреле число очков описывается тождественной случайной переменной f на множестве исходов $U = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$: $f(u) = u$ для каждого исхода u. Как нетрудно проверить,

$$f = \sum_{0 \leqslant u \leqslant 10} u \cdot i_u.$$

2.2*. Элементарное распределение

Каждая случайная переменная определяет некоторую положительную нормированную функцию на множестве своих значений. Эта функция называется ее *элементарным распределением*.

Рассмотрим случайную переменную f на множестве исходов U и ее множество значений

$$X = f(U).$$

Примеры. В примерах 1—3 из § 1 соответственно:

$X = \{+1, -1\}$; $X = \{+10, 0, -10\}$; $X = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$.

Для каждого числа x из множества X определено событие

$$f^{-1}(x) = \{u : f(u) = x\},$$

составленное из всех исходов u, для которых значение случайной переменной $f(u)$ равно числу x.

Примеры. В примерах 1—3 из § 1 соответственно:

$$f^{-1}(+1)=\{2, 4, 6\},\ f^{-1}(-1)=\{1, 3, 5\};$$
$$f^{-1}(+10)=\{11\},\ f^{-1}(0)=\{10, 01\},\ f^{-1}(-10)=\{0, 0\};$$
$$f^{-1}(u)=\{u\}.$$

Для каждого события $f^{-1}(x)$ определена вероятность

$$q(x)=P(f^{-1}(x))=P(\{u:f(u)=x\}).$$

Примеры. В примерах 1—3 из § 1 соответственно:

$$q(+1)=1/2,\ q(-1)=1/2;\ q(+10)=1/4,\ q(0)=1/2,$$
$$q(-10)=1/4;\ q(u)=p(u).$$

Рассмотрим функцию q на множестве X со значением

$$q(x)=P(f^{-1}(x))$$

для каждого числа $x\in X$.

Лемма 1. *q является положительной нормированной функцией на множестве X.*

Доказательство. *Положительность* q следует из положительности вероятности P.

Нормированность q следует из нормированности вероятности P. Действительно, события $f^{-1}(x)=\{u:f(u)=x\}$ семейства $\{f^{-1}(x)\}_{x\in X}$ попарно не пересекаются: если $u\in f^{-1}(x)\cap f^{-1}(y)$, то $x=f(u)=y$. Сумма $\Sigma f^{-1}(x)$ всех событий $f^{-1}(x)$ семейства $\{f^{-1}(x)\}_{x\in X}$ равна множеству исходов U:

$$\Sigma f^{-1}(x)=U.$$

По общему правилу сложения отсюда следует, что

$$\Sigma q(x)=\Sigma P(f^{-1}(x))=P(\Sigma f^{-1}(x))=P(U)=1.$$

Лемма 1 доказана.

Примеры. В примерах 1—3 из § 1 соответственно:

$$\Sigma q(x)=q(1)+q(-1)=P(\{2, 4, 6\})+P(\{1, 3, 5\})=1;$$
$$\Sigma q(x)=q(10)+q(0)+q(-10)=P(\{11\})+P(\{10, 01\})+$$
$$+P(\{00\})=1.$$
$$\sum q(x)=\sum_{u\in U} p(u)=\sum_{k=0}^{10} p(k)=1.$$

Лемма 1 позволяет сформулировать

Определение 2. *Элементарным распределением для случайной переменной f называется положительная нормированная функция q на множестве X значений случайной переменной, имеющая значение*

$$q(x)=P(f^{-1}(x))$$

для каждого числа x из множества X.

Часто элементарное распределение q определяется таблицей своих значений: в верхней строке ее записываются числа x, а в нижней под каждым числом x — значение $q(x)$ элементарного распределения q для x.

Примеры. В примерах 1—3 из § 1 соответственно:

x	$+1$	-1
$q(x)$	1/2	1/2

;

x	10	0	—10
$q(x)$	1/4	1/2	1/4

;

x	0	1	2	3	4	5	6	7	8	9	10
$q(x)$	1/100	1/100	1/100	1/100	1/100	1/20	1/20	1/20	1/10	1/5	1/2

.

В примерах 1—3 пункта 1 из § 2 соответственно:

1) каждый из индикаторов i_1, i_2, i_3, i_4, i_5, i_6 имеет одно и то же элементарное распределение q с таблицей значений

x	0	1
$q(x)$	1/6	5/6

;

2) каждый из индикаторов i_{11}, i_{10}, i_{01}, i_{00} имеет одно и то же элементарное распределение q с таблицей значений

x	1	0
$q(x)$	1/4	3/4

;

3) для каждого исхода u его индикатор i_u имеет элементарное распределение q_u с таблицей значений

x	1	0
$q_u(x)$	$p(u)$	$1-p(u)$

.

В частности,

x	1	0
$q_0(x)$	1/100	99/100

;

x	1	0
$q_5(x)$	1/20	19/20

;

x	1	0
$q_8(x)$	1/10	9/10

;

x	1	0
$q_9(x)$	1/5	4/5

;

x	1	0
$q_{10}(x)$	1/2	1/2

.

3.2*. Распределение

Элементарное распределение для каждой случайной переменной определяет некоторую положительную, нормированную и аддитивную функцию на классе всех частей своего множества значений. Эта функция называется ее *распределением.* Она определяет вероятностные свойства случайной переменной.

Рассмотрим случайную переменную f на множестве исходов U, ее множество значений $X=f(U)$, элементарное распределение q и алгебру, образованную классом

$$\mathcal{B}=f[\mathcal{A}]$$

всех частей множества X, с определенными для нее объединением, пересечением и дополнением.

Примеры. В примерах 1, 2 из § 1 соответственно:

$$\mathcal{B}=\{O,\ \{+1\},\ \{-1\},\ X\};$$

$$\mathcal{B}=\{O,\ \{10\},\ \{0\},\ \{-10\},\ \{10,\ 0\},\ \{10,\ -10\},\ \{0,\ -10\},\ X\};$$

$$\mathcal{B}=\mathcal{A}.$$

Для каждой части B множества X определена сумма

$$Q(B)=\sum_{x\in B} q(x)$$

значений $q(x)$ элементарного распределения q для чисел x из множества B.

Примеры. В примерах 1—3 из § 1 соответственно:

$$Q(\{1\})=q(1)=1/2;$$

$$Q(\{10,\ 0\})=q(10)+q(0)=3/4;$$

$$Q(\{6,\ 7,\ 8,\ 9,\ 10\})=q(6)+q(7)+q(8)+q(9)+q(10)=9/10.$$

Рассмотрим функцию Q на классе $\mathcal{B}$ со значением $Q(B)$ для каждого множества B из класса $\mathcal{B}$.

Примеры. В примерах 1—3 из § 1 соответственно:

$$Q(O)=0,\ Q(\{1\})=1/2,\ Q(\{-1\})=1/2,\ Q(X)=1;$$

B	O	$\{10\}$	$\{0\}$	$\{-10\}$	$\{10,0\}$	$\{10,\ -10\}$	$\{0,\ -10\}$	X
$Q(B)$	0	1/4	1/2	1/4	3/4	1/2	3/4	1

;

$$Q(B)=P(B).$$

Лемма 2. *Q является положительной, нормированной и аддитивной функцией на классе $\mathcal{B}$.*

Доказательство. По определениям 3 и 4 из § 2 главы 3 части I, функция q является *элементарной вероятностью,* а функция Q — *вероятностью* для вероятностной модели $(X,\ q;\ \mathcal{B},\ Q)$ с *множеством исходов* X и *алгеброй событий* $\mathcal{B}$. По теореме об основных

свойствах вероятность Q является положительной, нормированной и аддитивной функцией на классе $\mathscr{B}$. Лемма 2 доказана.

Для каждой части B множества значений X случайной переменной f определено событие

$$f^{-1}(B)=\{u:f(u)\in B\},$$

составленное из всех исходов u, для которых значение $f(u)$ случайной переменной f принадлежит множеству B.

Примеры. В примерах 1—3 из § 1 соответственно:

$$f^{-1}(O)=O,\ f^{-1}(\{1\})=\{2,\ 4\ 6\},\ f^{-1}(\{-1\})=\{1,\ 3,\ 5\},\ f^{-1}(X)=U;$$

B	O	$\{10\}$	$\{0\}$	$\{-10\}$	$\{10, 0\}$	$\{10, -10\}$	$\{0, -10\}$	X
$f^{-1}(B)$	O	$\{11\}$	$\{10, 01\}$	$\{00\}$	$\{11, 10, 01\}$	$\{11, 00\}$	$\{10, 01, 00\}$	U

;

$$f^{-1}(B)=B.$$

Для каждого события $f^{-1}(B)=\{u:f(u)\in B\}$ определена вероятность

$$Pf^{-1}(B)=P(f^{-1}(B))=P(\{u:f(u)\in B\}).$$

Примеры. В примерах 1—3 из § 1 соответственно:

$$Pf^{-1}(O)=0,\ Pf^{-1}(\{1\})=1/2,\ Pf^{-1}(\{-1\})=1/2,\ Pf^{-1}(X)=1;$$

B	O	$\{10\}$	$\{0\}$	$\{-10\}$	$\{10,0\}$	$\{10,-10\}$	$\{0, -10\}$	X
$Pf^{-1}(B)$	0	1/4	1/2	1/4	3/4	1/2	3/4	1

;

$$Pf^{-1}(B)=P(B).$$

Рассмотрим функцию Pf^{-1} на классе $\mathscr{B}$ со значением $Pf^{-1}(B)$ для каждого множества B из класса $\mathscr{B}$.

Лемма 3. $Q=Pf^{-1}$.

Доказательство. Для каждого множества B из класса $\mathscr{B}$ рассмотрим семейство $\{f^{-1}(x)\}_{x\in B}$ событий $f^{-1}(x)$. События $f^{-1}(x)$ попарно не пересекаются. Сумма $\sum\limits_{x\in B} f^{-1}(x)$ этих событий равна множеству $f^{-1}(B)$:

$$\sum_{x\in B} f^{-1}(x)=\sum_{x\in B}\{u:f(u)=x\}=\{u:f(u)\in B\}=f^{-1}(B).$$

По общему правилу сложения отсюда следует, что

$$Q(B)=\sum_{x\in B} q(x)=\sum_{x\in B} P(f^{-1}(x))=P\left(\sum_{x\in B} f^{-1}(x)\right)=P(f^{-1}(B))=$$
$$=Pf^{-1}(B).$$

Лемма 3 доказана.

Приведенные в примерах таблицы значений функций Q и Pf^{-1} поясняют равенство леммы 3. Ее доказательство иллюстрируют

Примеры. В примерах 1—3 из § 1 соответственно:

$$Q(\{1\}) = q(1) = P(f^{-1}(1)) = Pf^{-1}(\{1\});$$

$$Q(\{10, 0\}) = q(10) + q(0) = P(f^{-1}(10)) + P(f^{-1}(0)) =$$
$$= P(f^{-1}(10) + f^{-1}(0)) = P(f^{-1}(\{10, 0\}) = Pf^{-1}(\{10, 0\});$$

$$Q(B) = P(B) = Pf^{-1}(B).$$

Леммы 2 и 3 позволяют сформулировать

Определение 3. *Распределением для случайной переменной называется положительная, нормированная и аддитивная функция $Q = Pf^{-1}$ на классе $\mathcal{B}$ всех частей множества значений X случайной переменной f, имеющая значение*

$$Q(B) = P(f^{-1}(B))$$

для каждого множества B из класса $\mathcal{B}$.

Примеры. В примерах 1, 2 из § 1 значения распределений $Q = Pf^{-1}$ описываются рассматривавшимися выше равенствами и таблицей. В примере 3 из § 1 распределение Q равно вероятности P.

4.2*. Новая вероятностная модель

Вероятностная модель $(U, \mathcal{A}; p, P)$ и случайная переменная f на множестве исходов U определяют новую вероятностную модель $(X, \mathcal{B}; q, Q)$.

В этой новой модели:

1) *множество исходов* является множеством значений

$$X = f(U)$$

случайной переменной f;

2) *алгебра событий* образована классом

$$\mathcal{B} = f(\mathcal{A})$$

всех частей множества значений X случайной переменной f и определенными для этого класса объединением, пересечением и дополнением;

3) *элементарная вероятность* является элементарным распределением q случайной переменной f и имеет значение

$$q(x) = P(f^{-1}(x))$$

для каждого числа x из множества X;

4) *вероятность* является распределением Q случайной переменной f и имеет значение

$$Q(B) = P(f^{-1}(B))$$

для каждого множества B из класса $\mathscr{B}$.

Примеры. Примеры 1—3 из § 1.

Условимся элементарную вероятность q обозначать тем же символом Pf^{-1}, что и вероятность Q:

$$q = Pf^{-1}.$$

Это оправдывается тем, что значения q равны значениям Q на элементарных событиях:

$$q(u) = Pf^{-1}(u) = Pf^{-1}(\{u\}) = Q(\{u\})$$

для каждого исхода u из множества U.

5.2*. Механическая интерпретация

Вероятностная модель $(U, \mathscr{A}; p, P)$ представляет также механическую *систему точек* u из множества U, в каждой из которых расположена *масса* $p(u)$.

Случайная переменная f отображает множество U на множество X точек числовой прямой. При этом в точку x из множества X отображаются точки тела $f^{-1}(x)$. В точке x поэтому располагается масса $q(x)$, равная массе $P(f^{-1}(x))$ тела $f^{-1}(x)$. Новая вероятностная модель $(X, \mathscr{B}; q, Q)$ представляет также новую механическую систему точек x из множества X, в каждой из которых расположена масса $q(x)$.

Таким образом, случайная переменная f переносит массу $p(u)$ из точки u множества U в точку $x = f(u)$ числовой прямой. Общая масса $q(x)$, перенесенная в точку x, равна сумме всех масс $p(u)$, перенесенных в эту точку.

Примеры. В примерах 1, 2 из § 1 перенос массы поясняется соответственно рис. 21 и 22:

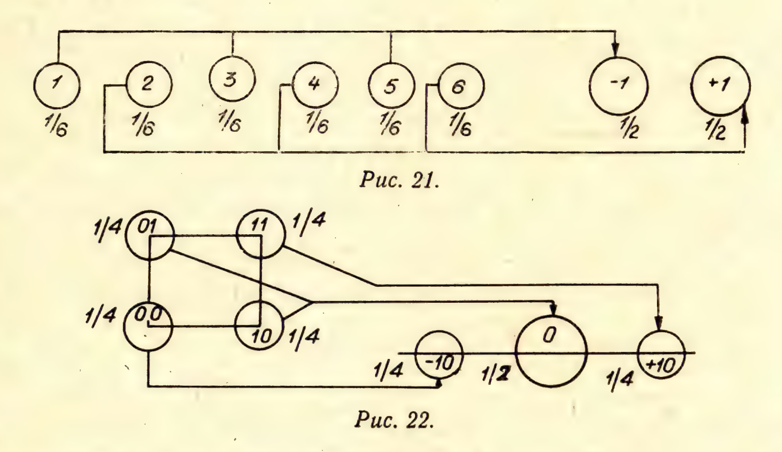


Рис. 21.

Рис. 22.

В примере 3 из § 1 масса остается на месте.

6.2*. Примеры распределений

Приведем несколько примеров часто встречающихся распределений.

1.6.2. Равномерное распределение

Рассмотрим стандартную модель Лапласа n равновероятных исходов 1, 2, ..., n и тождественную случайную переменную e на множестве исходов $u = \{1, 2, \ldots, n\}$:

$$e(n) = n$$

для каждого исхода n.

Элементарное распределение q для случайной переменной e равно элементарной вероятности p:

$$q(x) = p(u) = 1/n$$

для каждого значения $x = u$ случайной переменной e.

Распределение Q для случайной переменной e равно вероятности P:

$$Q(B) = P(A) = n(A)/n(U)$$

для каждой части $B = A$ множества значений $X = U$ случайной переменной e.

Распределение Q случайной переменной e называется *равномерным*: единица массы распределена равномерно по точкам 1, 2, ..., n.

2.6.2. Биномиальное распределение

Рассмотрим модель Бернулли n испытаний с вероятностью успеха a и случайную переменную s на множестве исходов $U = B^n$, значение $s(u)$ которой для каждой строки $u = u_1, \ldots, u_n$ равно сумме элементов строки:

$$s(u) = \sum_{k=1}^{n} u_k.$$

Случайная переменная s описывает число успехов при n испытаниях.

Примеры. Для $n = 1, 2, 3$ таблицы значений случайной переменной s имеют вид:

n	0	1
$s(n)$	0	1

;

n	00	01	10	11
$s(n)$	0	1	1	2

;

n	000	001	010	100	011	101	110	111
$s(n)$	0	1	1	1	2	2	2	3

.

В примере 1 пункта 4 из § 3 главы 3 части I было показано, что для каждого значения $m=0, 1, \ldots, n$ случайной переменной s верны равенства

$$q(m) = P(\{u : s(u) = m\}) = \binom{n}{m} a^m (1-a)^{n-m}.$$

Примеры. Для $n=1, 2, 3$ таблицы значений элементарного распределения q случайной переменной s имеют вид:

m	1	0
$q(m)$	a	$1-a$

;

m	2	1	0
$q(m)$	a^2	$2a(1-a)$	$(1-a)^2$

m	3	2	1	0
$q(m)$	a^3	$3a^2(1-a)$	$3a(1-a)^2$	$(1-a)^3$

.

Распределение Q случайной переменной s называется *биномиальным*. Для каждой части B множества $X=\{0, 1, \ldots, n\}$ значение биномиального распределения Q равно числу

$$Q(B) = \sum_{m \in B} \binom{n}{m} a^m (1-a)^{n-m}.$$

3.6.2. Вырожденное распределение

Рассмотрим вероятностную модель $(U, \mathscr{A}; p, P)$ и случайную переменную f на множестве исходов U, значение $f(u)$ которой для каждого исхода u равно 0:

$$f(U)=\{0\}.$$

Множество значений X случайной переменной f состоит из единственного числа 0:

$$X=f(U)=\{0\}.$$

Случайная переменная f равна постоянной 0. Элементарное распределение q для случайной переменной f имеет единственное значение:

$$q(0)=1.$$

Распределение Q для случайной переменной f имеет следующую таблицу значений:

B	$\varnothing$	X
$Q(B)$	0	1

.

Оно является примером *вырожденного* распределения: вся масса сосредоточена в одной точке.

Более общим образом рассмотрим произвольное число c и случайную переменную f на U, значение которой равно c для каждого исхода u, для которого $p(u)>0$:

$$f(u)=c \quad (p(u)>0).$$

Значения $f(u)$ случайной переменной f для исходов u, для которых $p(u)=0$, могут быть не равны c. Говорят, что случайная переменная f равна постоянной c *почти всюду*.

Множество X значений такой случайной переменной может быть произвольным конечным множеством чисел.

Элементарное распределение q для случайной переменной f имеет значения:

$$q(x)=\begin{cases}1 & (x=c),\\ 0 & (x\neq c).\end{cases}$$

Распределение Q для случайной переменной f имеет значения:

$$Q(B)=\begin{cases}1 & (c\in B),\\ 0 & (c\notin B).\end{cases}$$

Оно называется *вырожденным* распределением: вся масса сосредоточена в одной точке.

Примерами случайных переменных, имеющих вырожденные распределения, являются индикаторы успеха и неудачи для модели Бернулли $n=1$ испытания с вероятностью успеха $a=0$ и $a=1$.

§ 3. НЕЗАВИСИМОСТЬ И ЗАВИСИМОСТЬ

Каждая случайная переменная определяет некоторый *класс событий*, описывающих ее значения. Естественно независимость случайных переменных отождествлять с независимостью определяемых ими классов событий. Независимость случайных переменных, как и независимость событий, имеет *вероятностный* характер. Содержание этого понятия поясняется примерами пункта 1.

Во всех случаях, когда нет соответствующих уточнений, случайные переменные, рассматриваемые в этом параграфе, определены на множестве исходов U произвольной конечной вероятностной модели $(U, \mathcal{A}; p, P)$.

При первом чтении рекомендуется ограничиться пунктами 1 и 2.

1.3. Примеры

Рассмотрим несколько примеров независимых и зависимых случайных переменных.

1.1.3. Пример 1

Для модели Бернулли $n=2$ испытаний с вероятностью успеха a рассмотрим индикаторы $f_1 = i_{У_1}$, $f_2 = i_{У_2}$ событий $У_1=\{11, 10\}$, $У_2=\{11, 01\}$:

u	11	10	01	00
$f_1(u)$	1	1	0	0

;

u	11	10	01	00
$f_2(u)$	1	0	1	0

.

Индикатор $f_1 = i_{У_1}$ определяет события $У_1 = f_1^{-1}(1)$ и $H_1 = f_1^{-1}(0) = \{01, 00\}$, а индикатор $f_2 = i_{У_2}$ — события $У_2 = f_2^{-1}(1)$ и $H_2 = f_2^{-1}(0) = \{10, 00\}$. События $У_1$ и $У_2$, $У_1$ и H_2, H_1 и $У_2$, H_1 и H_2 независимы, т. е. классы событий

$$\mathscr{C}_1 = \{f_1^{-1}(1), f_1^{-1}(0)\} = \{У_1, H_1\} \text{ и } \mathscr{C}_2 = \{f_2^{-1}(1), f_2^{-1}(0)\} = \{У_2, H_2\}$$

независимы. Естественно поэтому считать, что случайные переменные $f_1 = i_{У_1}$ и $f_2 = i_{У_2}$ *независимы*.

Индикаторы $f_1 = i_{У_1}$ и $f_2 = 1 - f_1 = i_{H_1}$ определяют одни и те же события $У_1 = f_1^{-1}(1) = f_2^{-1}(0)$ и $H_1 = f_1^{-1}(0) = f_2^{-1}(1)$. Если $0<a<1$, то события $У_1$ и H_1 зависимы. Следовательно, классы событий $\mathscr{C}_1 = \{f_1^{-1}(1), f_1^{-1}(0)\}$ и $\mathscr{C}_2 = \{f_2^{-1}(1), f_2^{-1}(0)\}$ зависимы. Естественно поэтому считать, что случайные переменные $f_1 = i_{У_1}$ и $f_2 = 1 - f_1 = i_{H_1}$ *зависимы*.

2.1.3. Пример 2

Для модели Лапласа $n=4$ равновероятных исходов 1, 2, 3, 4, рассмотрим индикаторы $f_1 = i_{C_1}$, $f_2 = i_{C_2}$, $f_3 = i_{C_3}$ событий $C_1=\{1, 4\}$, $C_2=\{2, 4\}$, $C_3=\{3, 4\}$:

u	1	2	3	4
$f_1(u)$	1	0	0	1

;

u	1	2	3	4
$f_2(u)$	0	1	0	1

;

u	1	2	3	4
$f_3(u)$	0	0	1	1

.

События C_1, C_2, C_3 рассматривались в примере Бернштейна с тетраэдром. Каждый индикатор f_j определяет события $C_j = f_j^{-1}(1)$, $C_j' = f_j^{-1}(0)$ $(j = 1, 2, 3)$. Как нетрудно проверить, события C_j и C_k, C_j и C_k', C_j' и C_k, C_j' и C_k' независимы для каждых различных номеров j и k из 1, 2, 3, т. е. классы событий

$$\mathscr{C}_j = \{f_j^{-1}(1), f_j^{-1}(0)\} = \{C_j, C_j'\} \text{ и } \mathscr{C}_k = \{f_k^{-1}(1), f_k^{-1}(0)\} = \{C_k, C_k'\}$$

независимы. Естественно поэтому считать, что случайные переменные f_1 и f_2, f_1 и f_3, f_2 и f_3 независимы. Можно сказать, что случайные переменные f_1, f_2, f_3 *попарно независимы.*

В то же время события C_1, C_2, C_3 зависимы. Следовательно, классы событий $\mathscr{C}_1 = \{f_1^{-1}(0), f_1^{-1}(0)\} = \{C_1, C_1'\}$, $\mathscr{C}_2 = \{f_2^{-1}(1), f_2^{-1}(0)\} = \{C_2, C_2'\}$, $\mathscr{C}_3 = \{f_3^{-1}(1), f_3^{-1}(0)\} = \{C_3, C_3'\}$ зависимы. Естественно поэтому считать, что случайные переменные f_1, f_2, f_3 *зависимы.*

3.1.3. Пример 3

Для модели Бернулли n испытаний с вероятностью успеха a рассмотрим случайные переменные f_1, f_2 со значениями

$$f_1(u) = \sum_{1 \leqslant k \leqslant n} u_k, \quad f_2(u) = \prod_{1 \leqslant k < n} u_k$$

для каждой строки $u = u_1 \ldots u_n$.

Случайная переменная f_1 определяет события $f_1^{-1}(m) = \left\{u : \sum_{1 \leqslant k \leqslant n} u_k = m\right\} (m = 0, 1, \ldots, n)$, случайная переменная f_2 — события $f_2^{-1}(1) = \{1 \ldots 1\}$, $f_2^{-1}(0) = \{u : u \neq 1 \ldots 1\}$. Если $0 < a < 1$, то события $f_1^{-1}(0) = \{0 \ldots 0\}$ и $f_2^{-1}(1) = \{1 \ldots 1\}$ зависимы:

$$P\left(f_1^{-1}(0) \cap f_2^{-1}(1)\right) = 0 \neq (1-a)^n a^n = P\left(f_1^{-1}(0)\right) P\left(f_2^{-1}(2)\right).$$

Следовательно, классы событий $\mathscr{C}_1 = \{f_1^{-1}(0), f_1^{-1}(1), \ldots, f_1^{-1}(n)\}$ и $\mathscr{C}_2 = \{f_2^{-1}(1), f_2^{-1}(0)\}$ зависимы. Естественно поэтому считать, что случайные переменные f_1 и f_2 *зависимы.*

2.3. Независимость двух случайных переменных

Рассмотрим случайные переменные f_1, f_2 с множествами значений

$$X_1 = f_1(U), \quad X_2 = f_2(U).$$

Случайная переменная f_1 определяет класс

$$\mathscr{C}_1 = \{f_1^{-1}(x_1) : x_1 \in X_1\},$$

составленный из событий

$$f_1^{-1}(x_1) = \{u : f_1(u) = x_1\},$$

образованных исходами u, для которых случайная переменная f_1 имеет значение $f_1(u)$, равное числу x_1 из множества X_1. Аналогично, случайная переменная f_2 определяет класс событий

$$\mathscr{C}_2 = \{f_2^{-1}(x_2) : x_2 \in X_2\}.$$

Классы $\mathscr{C}_1$ и $\mathscr{C}_2$ называются *независимыми*, если и только если каждые событие C_1 класса $\mathscr{C}_1$ и событие C_2 класса $\mathscr{C}_2$ независимы:

$$P(C_1C_2)=P(C_1)P(C_2) \quad (C_1\in\mathscr{C}_1,\ C_2\in\mathscr{C}_2).$$

Это определение позволяет сформулировать

Определение 4′. *Случайные переменные f_1 и f_2 называются независимыми, если определяемые ими классы событий $\mathscr{C}_1$ и $\mathscr{C}_2$ независимы.*

Если классы событий $\mathscr{C}_1$ и $\mathscr{C}_2$ зависимы, то случайные переменные f_1 и f_2 называются *зависимыми.*

Примеры. В примерах 1 и 2 пункта 1 случайные переменные f_1 и f_2 *независимы.* В примере 3 пункта 1 они *зависимы.*

1.2.3. Лемма о независимости двух случайных переменных

Эквивалентное определение независимости двух случайных переменных содержит

Лемма 1′. *Случайные переменные f_1 и f_2 независимы, если и только если равенство*

$$(1') \quad P\left(f_1^{-1}(x_1)\, f_2^{-1}(x_2)\right) = P\left(f_1^{-1}(x_1)\right) P\left(f_2^{-1}(x_2)\right)$$

верно для каждых значений x_1 и x_2 случайных переменных f_1 и f_2.

Доказательство. По определению, классы событий $\mathscr{C}_1$ и $\mathscr{C}_2$, определяемые случайными переменными f_1 и f_2, независимы, если и только если произвольные событие $C_1 = f_1^{-1}(x_1)$ из класса $\mathscr{C}_1$ и событие $C_2 = f_2^{-1}(x_2)$ из класса $\mathscr{C}_2$ независимы. По определению 7 пункта 1 § 5 главы 3 части I независимость этих событий эквивалентна условию леммы.

Примеры. 1. В примере 1 пункта 1 верны равенства

$$P\left(f_1^{-1}(1)\, f_2^{-1}(1)\right) = P(У_1У_2) = P(У_1)P(У_2) = P\left(f_1^{-1}(1)\right)P\left(f_2^{-1}(1)\right),$$

$$P\left(f_1^{-1}(1)f_2^{-1}(0)\right) = P(У_1H_2) = P(У_1)\, P(H_2) = P\left(f_1^{-1}(1)\right)P\left(f_2^{-1}(0)\right),$$

$$P\left(f_1^{-1}(0)\, f_2^{-1}(1)\right) = P(H_1У_2) = P(H_1)P(У_2) = P\left(f_1^{-1}(0)\right)P\left(f_2^{-1}(1)\right),$$

$$P\left(f_1^{-1}(0)\, f_2^{-1}(0)\right) = P(H_1H_2) = P(H_1)\, P(H_2) = P\left(f_1^{-1}(0)\right)P\left(f_2^{-1}(0)\right).$$

Случайные переменные f_1 и f_2 *независимы.*

2. В примере 3 пункта 1 имеем:

$$P\left(f_1^{-1}(0)\, f_2^{-1}(1)\right) \neq P\left(f_1^{-1}(0)\right) P\left(f_2^{-1}(1)\right),$$

случайные переменные f_1 и f_2 *зависимы.*

2.2.3*. Теорема о независимости двух случайных переменных

Эквивалентное определение независимости двух случайных переменных содержит также

Теорема 1′. *Случайные переменные f_1 и f_2 независимы, если и только если равенство*

$$(2') \quad P\left(f_1^{-1}(B_1)\, f_2^{-1}(B_2)\right) = P\left(f_1^{-1}(B_1)\right) P\left(f_2^{-1}(B_2)\right)$$

верно для каждых множеств B_1 и B_2 значений случайных переменных f_1 и f_2.

Доказательство. Для каждого множества B_j значений случайной переменной f_j рассмотрим семейство $\left(f_j^{-1}(x_j)\right)_{x_j \in B_j}$ событий $f_j^{-1}(x_j)$ $(j = 1, 2)$. События $f_j^{-1}(x_j)$ и $f_j^{-1}(y_j)$ при различных значениях x_j и y_j случайной переменной f_j не пересекаются.

Кроме того,

$$\sum_{x_j \in B_j} f_j^{-1}(x_j) = \sum_{x_j \in B_j} \{u : f_j(u) = x_j\} = \{u : f_j(u) \in B_j\} =$$

$$= f_j^{-1}(B_j) \quad (j = 1, 2).$$

Следовательно, события $f_1^{-1}(x_1)\, f_2^{-1}(x_2)$ и $f_1^{-1}(y_1)\, f_2^{-1}(y_2)$ при различных парах x_1, x_2 и y_1, y_2 значений случайных переменных f_1 и f_2 не пересекаются. Кроме того,

$$f_1^{-1}(B_1)\, f_2^{-1}(B_2) = \sum_{x_1 \in B_1} f_1^{-1}(x_1) \sum_{x_2 \in B_2} f_2^{-1}(x_2) = \sum_{x_1 \in B_1} \sum_{x_2 \in B_2} f_1^{-1}(x_1) f_2^{-1}(x_2).$$

Отсюда, используя правило сложения и лемму 1′, получаем:

$$P\left(f_1^{-1}(B_1)\, f_2^{-1}(B_2)\right) = \sum_{x_1 \in B_1} \sum_{x_2 \in B_2} P\left(f_1^{-1}(x_1)\, f_2^{-1}(x_2)\right) =$$

$$= \sum_{x_1 \in B_1} \sum_{x_2 \in B_2} P\left(f_1^{-1}(x_1)\right) P\left(f_2^{-1}(x_2)\right) = \sum_{x_1 \in B_1} P\left(f_1^{-1}(x_1)\right) \sum_{x_2 \in B_2} P\left(f_2^{-1}(x_2)\right) =$$

$$= P\left(f_1^{-1}(B_1)\right) P\left(f_2^{-1}(B_2)\right).$$

Теорема 1′ доказана.

Пример. В примере 3 пункта 1 для $B_1 = \{0, \ldots, n-1\}$ и $B_2 = \{1\}$ имеем, если $0 < a < 1$:

$$P\left(f_1^{-1}(B_1)\, f_2^{-1}(B_2)\right) = P(0) \neq (1-a)^n a^n = P\left(f_1^{-1}(B_1)\right) P\left(f_2^{-1}(B_2)\right).$$

Случайные переменные f_1 и f_2 зависимы.

3.3*. Независимость семейства случайных переменных

Рассмотрим конечное семейство $\{f_j\}_{j \in I}$ случайных переменных f_j с множествами значений

$$X_j = f_j(U).$$

Каждая случайная переменная f_j определяет класс

$$\mathscr{C}_j = \{f_j^{-1}(x_j) : x_j \in X_j\},$$

составленный из событий

$$f_j^{-1}(x_j) = \{u : f_j(u) = x_j\},$$

образованных исходами u, для которых случайная переменная f_j имеет значение $f_j(u)$, равное числу x_j из множества X_j.

Определение 4. *Случайные переменные f_j семейства $(f_j)_{j\in I}$ называются независимыми, если определяемые ими классы событий $\mathscr{C}_j$ семейства $(\mathscr{C}_j)_{j\in I}$ независимы.*

Если классы событий $\mathscr{C}_j$ зависимы, то случайные переменные f_j называются *зависимыми.*

Замечание. Если $I = \{1, 2\}$, то определение 4 эквивалентно определению 4′ пункта 1.

Пример. Рассмотрим модель Бернулли n испытаний с вероятностью успеха a. Для каждого номера $j = 1, 2, \ldots n$ рассмотрим случайную переменную f_j со значением $f_j(u) = u_j$ для каждого исхода $u = u_1 \ldots u_n$. Каждая из случайных переменных f_j имеет множество значений

$$X_j = f_j(U) = \{1, 0\}$$

и определяет класс

$$\mathscr{C}_j = \{f_j^{-1}(1), \quad f_j^{-1}(0)\} = \{У_j, \ H_j\},$$

составленный из событий

$$f_j^{-1}(1) = \{u : u_j = 1\} = У_j; \ f_j^{-1}(0) = \{u : u_j = 0\} = H_j.$$

Из теоремы о независимости испытаний в модели Бернулли, доказанной в пункте 2 § 5 главы 2 части I, следует, что классы событий $\mathscr{C}_j$ семейства $(\mathscr{C}_j)_{1\leqslant j\leqslant n}$ независимы. Следовательно, случайные переменные f_j рассматриваемого семейства $(f_j)_{1\leqslant j\leqslant n}$ *независимы.*

Замечание. Случайные переменные f_j являются индикаторами событий $У_j$. Для модели Бернулли $n = 2$ испытаний они рассматривались в примере 1 пункта 1.

1.3.3. Лемма о независимости семейства случайных переменных

Эквивалентное определение независимости семейства случайных переменных содержит

Лемма 1. *Случайные переменные f_j семейства $(f_j)_{j\in I}$ независимы, если и только если равенство*

$$(1) \qquad P\Big(\bigcap_{j\in I} f_j^{-1}(x_j)\Big) = \prod_{j\in I} P(f_j^{-1}(x_j))$$

верно для каждого семейства $(x_j)_{j\in I}$ значений x_j случайных переменных f_j.

Доказательство. 1. Если случайные переменные f_j семейства $(f_j)_{j\in I}$ независимы, то равенство (1) следует из определения 4 независимости случайных переменных и определений 11 и 12 независимости событий и классов событий из § 5 главы 3 части I.

2. Предположим, что равенство (1) верно, и докажем, что в этом случае равенство

$$(2) \qquad P\left(\bigcap_{k\in K} f_k^{-1}(x_k)\right) = \prod_{k\in K} P\left(f_k^{-1}) x_k)\right)$$

верно для каждого множества $K\subseteq I$ и каждого семейства $(x_k)_{k\in K}$ значений x_k случайных переменных f_k с индексами k из множества K.

Заметим, что для каждого индекса j из множества I события $f_j^{-1}(x_j)$ класса $\mathscr{C}_j$ попарно не пересекаются и верны равенства:

$$\sum_{x_j\in X_j} f_j^{-1}(x_j) = U,$$

$$\sum_{x_j\in X_j} (P f_j^{-1}(x_j)) = 1.$$

Для индексов L из множества $L=I-K$ пересечение первых и произведение вторых равенств дают:

$$\sum\left(\bigcap_{l\in L} f_l^{-1}(x_l)\right) \bigcap_{l\in L}\left(\sum_{x_l\in X_l} f_l^{-1}(x_l)\right) = U,$$

$$\sum\left(\prod_{l\in L} P(f_l^{-1}(x_l))\right) = \prod_{l\in L}\left(\sum_{x_l\in X_l} P(f_l^{-1}(x_l))\right) = 1.$$

Отсюда следует, что

$$\bigcap_{k\in K} f_k^{-1}(x_k) = \bigcap_{k\in K} f_k^{-1}(x_k) \sum\left(\bigcap_{l\in L} f_l^{-1}(x_l)\right) =$$

$$= \sum\left(\bigcap_{k\in K} f_k^{-1}(x_k) \bigcap_{l\in L} f_k^{-1}(x_l) = \sum\left(\bigcap_{j\in I} f_j^{-1}(x_j)\right);\right.$$

$$\prod_{k\in K} P(f_k^{-1}(x_k)) = \prod_{k\in K} P(f_k^{-1}(x_k)) \sum\left(\prod_{l\in L} P(f_l^{-1}(x_l))\right) =$$

$$= \sum\left[\prod_{k\in K} P(f_k^{-1}(x_k)) \prod_{l\in L} P(f_l^{-1}(x_l))\right] = \sum\left(\prod_{j\in I} P(f_j^{-1}(x_j))\right).$$

Используя эти равенства, общее правило сложения и равенство (1), получаем:

$$(3)\ P\left(\bigcap_{k\in K} f_k^{-1}(x_k)\right) = P\left(\sum\left(\bigcap_{j\in I} f_j^{-1}(x_j)\right)\right) = \sum P\left(\bigcap_{j\in I} f_j^{-1}(x_j)\right) =$$

$$= \sum\left[\prod_{j\in I} P(f_j^{-1}(x_j))\right] = \prod_{k\in K} P(f_k^{-1}(x_k)).$$

По определениям 11 и 12 независимости событий и классов событий из § 5 главы 3 части I и определения 4 независимости случайных переменных из равенства (3) следует, что случайные переменные f_j семейства $(f_j)_{j\in I}$ независимы.

Лемма 1 доказана.

Замечание. Если $I=\{1, 2\}$, то лемма 1 эквивалентна лемме 1' пункта 1. Использовавшиеся при доказательстве произведения для множества индексов L равны 1.

Пример. Рассмотрим множество $I=\{1, 2, \ldots, n\}$ номеров и семейство $(f_j)_{j\in I}$ индикаторов f_j событий $У_j$ для модели Бернулли n испытаний с вероятностью успеха a. Обозначим буквами k и l количества чисел $x_j=1$ и $x_j=0$ в произвольном семействе $(x_j)_{j\in I}$ значений x_j случайных переменных f_j. По формуле успехов и неудач из § 2 главы 2 части I верны равенства

$$P\left(\bigcap_{j\in I} f_j^{-1}(x_j)\right) = a^k(1-a)^l = \prod_{j\in I} P\left(f_j^{-1}(x_j)\right).$$

Равенство (1) верно, и случайные переменные f_j семейства $(f_j)_{j\in I}$ *независимы.*

2.3.3. Теорема о независимости семейства случайных переменных

Эквивалентное определение независимости семейства случайных переменных содержит также

Теорема 1. *Случайные переменные f_j семейства $(f_j)_{j\in I}$ независимы, если и только если равенство*

$$(2)\qquad P\left(\bigcap_{j\in I} f_j^{-1}(B_j)\right) = \prod_{j\in I} P\left(f_j^{-1}(B_j)\right)$$

верно для каждого семейства $(B_j)_{j\in I}$ множеств B_j значений случайных переменных f_j.

Доказательство. Для каждого множества B_j значений случайной переменной f_j рассмотрим семейство $\left(f_j^{-1}(x_j)\right)_{x_j\in B_j}$ событий $f_j^{-1}(x_j)$ $(j\in I)$. События $f_j^{-1}(x_j)$ и $f_j^{-1}(y_j)$ при различных значениях x_j и y_j случайной переменной f_j не пересекаются. Кроме того,

$$\sum_{x_j\in B_j} f_j^{-1}(x_j) = \sum_{x_j\in B_j}\{u: f_j(u)=x_j\} = \{u: f_j(u)\in B_j\} =$$
$$= f_j^{-1}(B_j) \quad (j\in I).$$

Следовательно, события $\bigcap_{j\in I} f_j^{-1}(x_j)$ и $\bigcap_{j\in I} f_j^{-1}(y_j)$ при различных семействах $(x_j)_{j\in I}$ и $(y_j)_{j\in I}$ значений x_j и y_j случайных переменных f_j не пересекаются. Кроме того,

$$\bigcap_{j\in I} f_j^{-1}(B_j) = \bigcap_{j\in I}\sum_{x_j\in B_j} f_j^{-1}(x_j) = \sum \bigcap_{j\in I} f_j^{-1}(x_j).$$

Отсюда, используя общее правило сложения и лемму 1, получаем:

$$P\left(\bigcap_{j\in I} f_j^{-1}(B_j)\right) = \sum P\left(\bigcap_{j\in I} f_j^{-1}(x_j)\right) = \sum \prod_{j\in I} P\left(f_j^{-1}(x_j)\right) =$$
$$= \prod_{j\in I}\sum_{x_j\in B_j} P\left(f_j^{-1}(x_j)\right) = \prod_{j\in I} P\left(f_j^{-1}(B_j)\right).$$

Теорема 1 доказана.

Замечание. Если $I=\{1, 2\}$, то теорема 1 эквивалентна теореме 1' пункта 1.

Пример. Рассмотрим модель Бернулли n испытаний с вероятностью успеха a. Для каждого номера j из множества $I=\{1, \ldots, n\}$ рассмотрим произвольное событие A_j из испытания $\mathcal{B}_j = \{O, \mathcal{Y}_j, H_j, U\}$ и индикатор $f_j = i_{A_j}$ события A_j. Для каждого исхода v, принадлежащего событию A_j, значение индикатора $f_j = i_{A_j}$ равно 1, а для каждого исхода v, не принадлежащего событию A_j, значение индикатора $f_j = i_{A_j}$ равно 0:

$$f_j(v) = i_{A_j}(v) = \begin{cases} 1 & (v \in A_j), \\ 0 & (v \notin A_j). \end{cases}$$

Заметим, что

$$i_O(U) = \{0\}, \quad i_{\mathcal{Y}_j}(U) = i_{H_j}(U) = \{1, 0\}, \quad i_U(U) = \{1\};$$

$$i_O^{-1}(\{0\}) = U, \quad i_{\mathcal{Y}_j}^{-1}(\{1\}) = i_{H_j}^{-1}(\{0\}) = \mathcal{Y}_j,$$

$$i_{\mathcal{Y}_j}^{-1}(\{0\}) = i_{H_j}^{-1}(\{1\}) = H_j, i_U^{-1}(\{1\}) = U.$$

Из этих равенств следует, что каждое множество B_j значений случайной переменной $f_j = i_{A_j}$ определяет исход $f_j^{-1}(B_j) = i_{A_j}^{-1}(B_j) = C_j$ испытания $\mathcal{B}_j$. По лемме пункта 1 § 5 главы 2 части I исходы C_j независимы и

$$P\left(\bigcap_{j\in I} f_j^{-1}(B_j)\right) = P\left(\bigcap_{j\in I} C_j\right) = \prod_{j\in I} P(C_j) = \prod_{j\in I} P\left(f_j^{-1}(B_j)\right).$$

Равенство (2) верно, и случайные переменные f_j семейства $(f_j)_{j\in I}$ независимы.

4.3*. Попарная независимость

Часто рассматривается лишь попарная независимость случайных переменных.

Определение 5. *Случайные переменные f_j семейства $(f_j)_{j\in I}$ называются попарно независимыми, если и только если для каждых двух различных индексов j и k случайные переменные f_j и f_k независимы.*

Замечание. Если случайные переменные f_j семейства $(f_j)_{j\in I}$ независимы, то они попарно независимы. Это следует из сравнения определений 4 и 5. Обратное предположение неверно.

Пример. В примере 2 пункта 1 случайные переменные f_j семейства $(f_j)_{j\in I}$ с множеством индексов $I=\{1, 2, 3\}$ зависимы. В то же время они попарно независимы.

5.3*. Совместное распределение

Понятие независимости случайных переменных можно пояснить, используя понятие их совместного распределения.

1.5.3. Примеры

Рассмотрим несколько примеров совместных распределений.

Пример 1. Для модели Бернулли $n=2$ испытаний с вероятностью успеха a рассмотрим пару $f=(f_1, f_2)$ индикаторов $f_1=i_{У_1}$ и $f_2=i_{H_2}$ событий $У_2=\{11, 10\}$ и $H_2=\{10, 00\}$. Случайные переменные f_1 и f_2 независимы.

Каждая из случайных переменных f_1, f_2 в отдельности определяет новую вероятностную модель

$$(X_1, \mathscr{B}_1; q_1, Q), \ (X_2, \mathscr{B}_2; q_2, Q_2).$$

Множества исходов X_1, X_2 и таблицы значений элементарных вероятностей q_1, q_2 для этих моделей имеют вид:

$$X_1=X_2=\{1, 0\},$$

x_1	1	0
$q_1(x_1)$	$P(У_1)=a$	$P(H_1)=1-a$

;

x_2	1	0
$q_2(x_2)$	$P(H_2)=1-a$	$P(У_2)=a$

.

Пара $f=(f_1, f_2)$ тоже определяет новую вероятностную модель $(X, \mathscr{B}; q, Q)$.

В этой модели:

1) *множество исходов*

$$X=f[U]=f_1[U]\times f_2[U]=X_1\times X_2;$$

2) *алгебра событий* образована классом

$$\mathscr{B}=f[\mathscr{A}]=f_1[\mathscr{A}]\times f_2[\mathscr{A}]=\mathscr{B}_1\times\mathscr{B}_2$$

всех частей множества X и их объединением, пересечением и дополнением;

3) *элементарная вероятность* q имеет значения

$$q(x)=P(f^{-1}(x))=P(\{u: f(u)=x\})=$$
$$=P(\{u: f_1(u)=x_1, \ f_2(u)=x_2\})=$$
$$=P(\{u: f_1(u)=x_1\}\{u: f_2(u)=x_2\})=P\left(f_1^{-1}(x_1)f_2^{-1}(x_2)\right)$$

для каждого $x=(x_1, x_2)$ из множества X. Она называется совместным элементарным распределением случайных переменных f_1 и f_2;

4) *вероятность* Q имеет значение

$$Q(B)=P(f^{-1}(B))=P(\{u: f(u)\in B\}=$$
$$=P(\{u:(f_1(u), \ f_2(u))\in B\})=\sum_{x\in B} q(x)$$

для каждого B из класса $\mathscr{B}$. Она называется совместным распределением случайных переменных f_1 и f_2. Т. е.

$$X=\{11, 10, 01, 00\};$$

$$\mathscr{B}=\{O, \{11\}, \{10\}, \{01\}, \{00\}, \{11, 10\}, \{11, 01\},$$

$$\{11, 00\}, \{10, 01\}, \{10, 00\}, \{01, 00\}, \{11, 10, 01\},$$
$$\{11, 10, 00\}, \{11, 01, 00\}, \{10, 01, 00\}, X\};$$

x	11	10	01	00
$q(x)$	$P(У_1 Н_2) = a(1-a)$	$P(У_1 У_2) = a^2$	$P(Н_1 Н_2) = (1-a)^2$	$P(Н_1 У_2) = (1-a)a$

;

B	O	$\{11\}$	$\cdots$	$\{11, 10\}$	$\cdots$	$\{11, 10, 01\}$	$\cdots$	X
$Q(B)$	0	$q(11)$	$\cdots$	$q(11)+q(10)$	$\cdots$	$q(11)+q(10)+q(01)$	$\cdots$	1

.

Сравнивая таблицы значений элементарных распределений q_1, q_2 и q, замечаем, что вследствие независимости случайных переменных для каждого $x = (x_1, x_2)$ из множества $X = X_1 \times X_2$ верно равенство $q(x) = q_1(x_1) q_2(x_2)$:

$$P(У_1 У_2) = P(У_1) P(У_2), \ P(У_1 Н_2) = P(У_1) P(Н_2),$$
$$P(Н_1 Н_2) = P(Н_1) P(Н_2), \ P(Н_1 У_2) = P(Н_1) P(У_2).$$

Можно сказать, что совместное элементарное распределение независимых случайных переменных f_1 и f_2 равно произведению их элементарных распределений q_1 и q_2:

$$q = q_1 \times q_2.$$

Аналогично, для каждого прямоугольника $B = B_1 \times B_2$ из класса $\mathscr{B}$, равного декартову произведению множеств B_1 и B_2 из классов $\mathscr{B}_1$ и $\mathscr{B}_2$, верно равенство

$$Q(B) = Q_1(B_1) \times Q_2(B_2).$$

Если хотя бы одно из множеств B_1, B_2 пусто или равно X_1, X_2, то рассматриваемое равенство верно. Кроме того,

$$Q(\{1\} \times \{1\}) = Q_1(\{1\}) \times Q_2(\{1\}) = q(11),$$
$$Q(\{1\}) \times \{0\}) = Q(\{1\}) \times Q(\{0\} = q(10),$$
$$Q(\{0\} \times \{1\}) = Q_1(\{0\} \times Q_2(\{1\}) = q(01),$$
$$Q(\{0\}) \times \{0\}) = Q_1(\{0\}) \times Q(\{0\}) = q(00).$$

Можно сказать, что совместное распределение Q независимых случайных переменных f_1 и f_2 равно произведению их распределений Q_1 и Q_2:

$$Q = Q_1 \times Q_2.$$

Имея это в виду, можно сказать, что для независимых случайных переменных f_1 и f_2 *верно правило прямоугольника*.

В итоге можно сказать, что вероятностная модель, определяемая совместно независимыми случайными переменными f_1 и f_2, равна *произведению* вероятностных моделей, определяемых каждой из них в отдельности:

$$(X, \mathscr{B}; q, Q) = (X_1, \mathscr{B}_1; q_1, Q_1) \times (X_2, \mathscr{B}_2; q_2, Q_2).$$

Или, что то же самое,

$$X = X_1 \times X_2, \quad \mathscr{B} = \mathscr{B}_1 \times \mathscr{B}_2, \quad q = q_1 \times q_2, \quad Q = Q_1 \times Q_2.$$

Замечание. Для каждого из этих равенств умножение «$\times$» определяется по-своему. Однако введение различных обозначений уменьшило бы их общую выразительность.

Если вместо вероятностных моделей рассматривать механические системы, то сказанное можно пояснить рис. 23:

Рис. 23.

Система, определяемая совместно независимыми случайными переменными f_1 и f_2, состоит из точек 11, 10, 01, 00 плоскости, в которых расположены массы $a(1-a)$, a^2, $(1-a)^2$, $(1-a)a$. Системы, определяемые каждой из случайных переменных f_1 и f_2, состоят из точек 1, 0 на осях 1 и 2, в которых расположены массы a, $1-a$ и $1-a$, a. Независимость случайных переменных f_1 и f_2 выражается в том, что масса $q(x)$ каждой точки $x = (x_1, x_2)$ на плоскости *равна произведению масс* $q_1(x_1)$ и $q_2(x_2)$ масс точек x_1 и x_2 на осях 1 и 2.

Пример 2. Для модели Бернулли $n=2$ испытаний с вероятностью успеха a рассмотрим пару $f = (f_1, f_2)$ индикаторов $f_1 = i_{У_1}$ и $f_2 = i_{H_1}$ событий $У_1 = \{11, 10\}$ и $H_1 = \{01, 00\}$. Случайные переменные f_1 и f_2 зависимы.

Каждая из случайных переменных f_1, f_2 в отдельности определяет ту же вероятностную модель $(X_1, \mathscr{B}_1; q_1, Q_1)$, $(X_2, \mathscr{B}_2; q_2, Q_2)$, что и в примере 1.

По аналогии с примером 1 рассмотрим вероятностную модель $(X, \mathscr{B}; q, Q)$, определяемую совместно случайными переменными f_1, f_2.

Множество исходов и алгебра событий для этой модели те же, что и в примере 1. Таблица значений элементарной вероятности имеет вид

x	11	10	01	00
$q(x)$	$P(У_1 H_1) = 0$	$P(У_1 У_1) = a$	$P(H_1 H_1) = 1-a$	$P(H_1 У_1) = 0$

Если $0 < a < 1$, то

$$q(11) \neq q_1(1) q_2(1), \; q(10) \neq q_1(1) q_2(0), \; q(01) \neq q_1(0) q_2(1),$$
$$q(00) \neq q_1(0) q_2(0).$$

Совместное элементарное распределение q зависимых случайных переменных f_1 и f_2 не равно произведению их элементарных распределений q_1 и q_2:

$$q \neq q_1 \times q_2.$$

И следовательно, совместное распределение Q зависимых случайных переменных f_1 и f_2 не равно произведению их распределений Q_1 и Q_2:

$$Q \neq Q_1 \times Q_2.$$

Для зависимых переменных f_1 и f_2 ***не верно правило прямоугольника.***

Если вместо вероятностных моделей рассматривать механические системы, то сказанное можно пояснить рис. 24. Зависимость случайных переменных f_1 и f_2 выражается в том, что массы $q(x)$ некоторых точек $x=(x_1, x_2)$ на плоскости *не равны произведениям масс* $q_1(x_1)$ и $q_2(x_2)$ масс точек x_1 и x_2 на осях 1 и 2.

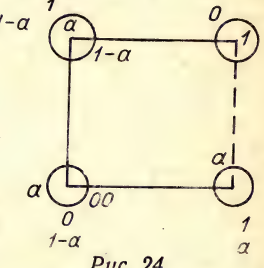


Рис. 24.

2.5.3. Определения

Рассмотрим семейство $f=(f_j)_{j \in I}$ случайных переменных f_j, имеющее конечное множество индексов I.

1.2.5.3. Произведение множеств значений. Для каждого исхода u семейство $f=(f_j)_{j \in I}$ случайных переменных f_j определяет семейство $f(u)=(f_j(u))_{j \in I}$ чисел $f_j(u)$, являющихся значениями случайных переменных f_j для исхода u.

Пример. Если $I=\{1, 2\}$, то $f=(f_1, f_2)$ и $f(u)=(f_1(u), f_2(u))$. В частности, в примере 1

$$f(11)=(i_{У_1}(11), i_{Н_2}(11))=(1, 0), \; f(10)=(i_{У_1}(10), i_{Н_2}(10))=(1, 1),$$
$$f(01)=(i_{У_1}(01), i_{Н_2}(01))=(0, 0), \; f(00)=(i_{У_1}(00), i_{Н_2}(00))=(0, 1).$$

Таким образом, семейство $f=(f)_{j \in I}$ отображает множество исходов U на декартово произведение

$$X=\prod_{j \in I} X_j$$

множеств значений X_j случайных переменных f_i.

Пример. Если $I=\{1, 2\}$, то $X=X_1 \times X_2$. В частности, в примере 1

$$X=\{1, 0\} \times \{1, 0\}.$$

2.2.5.3. Произведение алгебр. Рассмотрим алгебру, образованную классом

$$\mathscr{B}=f[\mathscr{A}]$$

всех частей множества X с определенными для них объединением, пересечением и дополнением. Среди различных множеств B класса $\mathscr{B}$ выделяются прямоугольники

$$B=\prod_{j \in I} B_j$$

со сторонами $B_j \subseteq X_j$ для каждого индекса j.

Пример. Если $I=\{1, 2\}$, то

$$B=B_1\times B_2.$$

В частности, в примере 1 прямоугольниками являются следующие множества:

$$O=O\times\{1, 0\}=\{1, 0\}\times O, \quad \{0\}\times\{1\},$$
$$\{1\}\times\{1\}, \ X=\{1, 0\}\times\{1, 0\}.$$

Условимся алгебру $\mathscr{B}$ называть *произведением алгебр* $\mathscr{B}_j=f[\mathscr{A}_j]$ и писать

$$\mathscr{B}=\prod_{j\in I}\mathscr{B}_j.$$

3.2.5.3. Произведение элементарных распределений. Для каждого семейства $x=(x_j)_{i\in I}$ из множества X определено событие

$$f^{-1}(x)=\{u:f(u)=x\}=\bigcap_{j\in I}\{u:f_j(u)=x_j\}=\bigcap_{j\in I}f_j^{-1}(x_j).$$

Пример. Если $I=\{1, 2\}$, то $x=(x_1, x_2)$ и

$$f^{-1}(x)=f_1^{-1}(x_1)f_2^{-1}(x_2).$$

В частности, в примере 1

$$f^{-1}(\{1, 1\})=У_1H_2, \ f^{-1}(\{1, 0\})=У_1У_2, \ f^{-1}(\{0, 1\})=H_1H_2,$$
$$f^{-1}(\{0, 0\})=H_1У_2.$$

Для каждого события $f^{-1}(x)$ определена вероятность

$$q(x)=P(f^{-1}(x))=P(\{u:f(u)=x\}).$$

Тем самым определена функция q на множестве X со значением $q(x)=P(f^{-1}(x))$ для каждого $x\in X$.

Лемма 1. *q является положительной нормированной функцией на множестве X.*

Доказательство дословно повторяет доказательство леммы 1 пункта 2 § 2. Лемма 1 позволяет сформулировать

Определение 2″. *Совместным элементарным распределением для случайных переменных f_j семейства $f=(f_j)_{j\in I}$ называется положительная нормированная функция q на множестве X семейств $x=(x_j)_{j\in I}$ значений случайных переменных f, имеющая значение*

$$q(x)=P\left(f^{-1}(x)\right)=P\left(\bigcap_{j\in I}f_j^{-1}(x_j)\right)$$

для каждого семейства $x=(x_j)_{j\in I}$ из множества X.

Замечание. Если $I=\{1\}$, то определение 2″ эквивалентно определению 2 пункта 2 § 2.

Условимся называть совместное элементарное распределение q произведением элементарных распределений q_j случайных переменных f_j семейства $f=(f_j)_{j\in I}$ и писать

$$q=\prod_{j\in I}q_j,$$

если и только если равенство

$$q(x) = \prod_{j \in I} q_j(x)$$

верно для каждого семейства $x = (x_j)_{j \in I}$ значений x_j случайных переменных f_j.

Примеры. Если $I = \{1, 2\}$, то $x = (x_1, x_2)$ и

$$q(x) = P\left(f_1^{-1}(x_1) f_2^{-1}(x_2)\right).$$

В частности, в примерах 1 и 2 приведены таблицы значений для совместного элементарного распределения q случайных переменных f_1 и f_2. В примере 1 совместное элементарное распределение q равно произведению элементарных распределений: $q = q_1 \times q_2$, а в примере 2 — не равно.

4.2.5.3. Произведение распределений. Для каждой части B множества X определена сумма

$$Q(B) = \sum_{x \in B} q(x)$$

значений $q(x)$ совместного элементарного распределения q для семейств x из множества B. Тем самым определена функция Q на классе $\mathscr{B}$ со значением $Q(B)$ для каждого множества B из класса $\mathscr{B}$.

Пример. В примере 1 приведена таблица значений для функции Q.

Лемма 2. *Q является положительной, нормированной и аддитивной функцией на классе $\mathscr{B}$*:

Доказательство повторяет доказательство леммы 2 пункта 2 § 2.

Для каждой части B множества X определено событие

$$f^{-1}(B) = \{u : f(u) \in B\},$$

составленное из всех исходов u, для которых семейство $f(u)$ принадлежит множеству B. Для каждого события $f^{-1}(B)$ определена вероятность

$$Pf^{-1}(B) = P(f^{-1}(B)) = P(\{u : f(u) \in B\}).$$

Тем самым определена функция Pf^{-1} на классе $\mathscr{B}$ со значением $Pf^{-1}(B)$ для каждого множества B из класса $\mathscr{B}$. Заметим, что

$$f^{-1}(B) = \{u : f(u) \in B\} = \bigcap_{j \in I} \{u : f_j(u) \in B_j\} = \bigcap_{j \in I} f_j^{-1}(B_j),$$

$$P\left(f^{-1}(B)\right) = P\left(\bigcap_{j \in I} f_j^{-1}(B_j)\right)$$

для каждого прямоугольника $B = \prod_{j \in I} B_j$ из класса $\mathscr{B}$.

Пример. Если $I = \{1, 2\}$, то $B = B_1 \times B_2$ и

$$P\left(f^{-1}(B)\right) = P\left(f_1^{-1}(B_1) f_2^{-1}(B_2)\right).$$

В частности, в примере 1

$$P\left(f^{-1}(\{1\}\times\{0\})\right)=P\left(f_1^{-1}(\{1\})\right)P\left(f_2^{-1}(\{0\})\right),$$

$$P\left(f^{-1}(\{0\})\times(\{1\})\right)=P\left(f_1^{-1}(\{0\})\times P\left(f_2^{-1}(\{1\})\right)\right).$$

Лемма 3″. $Q=Pf^{-1}$.

Доказательство повторяет доказательство леммы 3 пункта 2 § 2.

Леммы 2″ и 3″ позволяют сформулировать

Определение 3″. *Совместным распределением для случайных переменных* f_j *семейства* $(f_j)_{j\in I}$ *называется положительная, нормированная и аддитивная функция* $Q=Pf^{-1}$ *на классе* $\mathcal{B}$ *всех частей множества* X, *имеющая значение*

$$Q(B)=P\left(f^{-1}(B)\right)=P\left(\bigcap_{j\in I}f_j^{-1}(B_j)\right)$$

для каждого прямоугольника $B=\prod_{j\ni 1}B_j$ *из класса* $\mathcal{B}$.

Замечание. Если $I=\{1\}$, то определение 3″ эквивалентно определению 3 пункта 2 из § 2.

Условимся называть совместное распределение Q *произведением распределений* Q_j случайных переменных f_j семейства $(f_j)_{j\in I}$ и писать

$$Q=\prod_{j\in I}Q_j,$$

если и только если равенство

$$Q(B)=\prod_{j\in I}Q_j(B_j)$$

верно для каждого прямоугольника $B=\prod_{j\in I}B_j$, образованного из множеств B_j значений случайных переменных f_j.

Пример. В примере 1 приведена таблица значений совместного распределения Q случайных переменных f_1 и f_2. Оно равно произведению распределений Q_1 и Q_2 этих случайных переменных: $Q=Q_1\times Q_2$.

5.2.5.3. Произведение вероятностных моделей. Таким образом, случайные переменные f_j семейства $f=(f_j)_{j\in I}$ совместно определяют новую вероятностную модель $(X, \mathcal{B}; q, Q)$. В этой модели:

1) *множеством исходов* является множество X семейств $f(u)=(f_j(u))_{j\in I}$ значений $f_j(u)$ случайных переменных f_j;

2) *алгебра событий* образована классом $\mathcal{B}$ всех частей множества X и определенными для них объединением, пересечением и дополнением;

3) *элементарной вероятностью* является совместное элементарное распределение q случайных переменных f_j;

4) *вероятностью* является совместное распределение Q случайных переменных f_j.

Условимся называть вероятностную модель $(X, \mathscr{B}; q, Q)$ *произведением вероятностных моделей* $(X_j, \mathscr{B}_j; q_j, Q_j)$ и писать

$$(X, \mathscr{B}; q, Q) = \Pi(X_j, \mathscr{B}_j; q_j, Q_j),$$

если и только если

$$X = \Pi X_j, \ \mathscr{B} = \Pi \mathscr{B}_j, \ q = \Pi q_j, \ Q = \Pi Q_j \ (j \in I).$$

Примеры. В примере 1 вероятностная модель $(X, \mathscr{B}; q, Q)$ равна произведению вероятностных моделей $(X_1, \mathscr{B}_1; q_1, Q_1)$ и $(X_2, \mathscr{B}_2; q_2, Q_2)$, а в примере 2 — не равна.

3.5.3. Независимость

На языке распределений независимость случайных переменных означает, что их *совместное распределение равно произведению распределений* каждой из них.

Рассмотрим семейство $f = (f_j)_{j \in I}$ случайных переменных f_j, имеющее *конечное* множество индексов I. Для каждого индекса $j \in I$ случайная переменная f_j в отдельности определяет вероятностную модель $(X_j, \mathscr{B}_j; q_j, Q_j)$. Совместно случайные переменные f_j семейства $f = (f_j)_{j \in I}$ определяют вероятностную модель $(X, \mathscr{B}; q, Q)$.

Из определения 2″ и леммы 1 следует эквивалентная ей

Лемма 1″. *Случайные переменные f_j семейства $f = (f_j)_{j \in I}$ независимы, если и только если их совместное элементарное распределение q равно произведению элементарных распределений q_j случайных переменных f_j:*

$$q = \prod_{j \in I} q_j.$$

Из определения 3″ и теоремы 1 следует эквивалентная ей

Теорема 1″. *Случайные переменные f_j семейства $f = (f_j)_{j \in I}$ независимы, если и только если их совместное распределение Q равно произведению распределений Q_j случайных переменных f_j:*

$$Q = \prod_{j \in I} Q_j. \tag{2″}$$

Для равенства (2″) можно использовать также следующую выразительную форму:

$$Pf^{-1} = \prod_{j \in I} Pf_j^{-1}.$$

Лемму 1″ и теорему 1″ объединяет следующее предложение: случайные переменные f_j семейства $f = (f_j)_{j \in I}$ независимы, если и только если вероятностная модель $(X, \mathscr{B}; q, Q)$, определяемая ими совместно, равна произведению вероятностных моделей $X_j, \mathscr{B}_j; q_j, Q_j)$, определяемых каждой из случайных переменных f_j в отдельности:

$$(X, \mathscr{B}; q, Q) = \prod_{j \in I} (X_j, \mathscr{B}_j; q_j, Q_j).$$

Пример 3. Для модели Бернулли $n=3$ испытаний с вероятностью успеха $a=1/2$ рассмотрим тройку $f=(f_1, f_2, f_3)$ индикаторов $f_1=i_{У_1}$, $f_2=i_{У_2}$, $f_3=i_{У_3}$ событий $У_1$, $У_2$, $У_3$. Случайные переменные f_1, f_2, f_3 независимы. Их совместное элементарное распределение q равно произведению элементарных распределений q_1, q_2, q_3 каждой из них:

$$q=q_1\times q_2\times q_3.$$

Случайные переменные f_1, f_2, f_3 имеют одинаковые элементарные распределения $q_1=q_2=q_3$ с таблицей значений

x_j	1	0
$q_j(x_j)$	1/2	1/2

Таблица значений для совместного распределения случайных переменных f_1, f_2, f_3 имеет вид:

x	111	110	101	011	100	010	001	000
$q(x)$	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8

Если вместо вероятностных моделей рассматривать механические системы, то сказанное можно пояснить рис. 25.

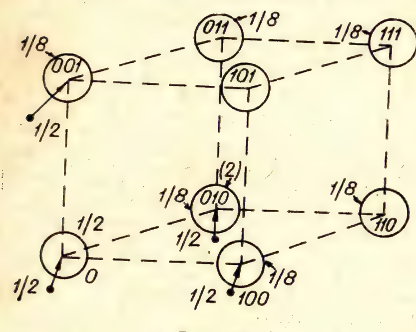


Рис. 25.

Независимость случайных переменных f_1, f_2, f_3 выражается в том, что масса $q(x)=1/8$ каждой точки $x=(x_1, x_2, x_3)$ в пространстве *равна произведению масс* $q_1(x_1)=q_2(x_2)=q_3(x_3)=1/2$ точек x_1, x_2, x_3 на осях 1, 2, 3.

Пример 4. Для модели Лапласа $n=4$ равновероятных исходов 1, 2, 3, 4 рассмотрим тройку $f=(f_1, f_2, f_3)$ индикаторов $f_1=i_{C_1}$, $f_2=i_{C_2}$, $f_3=i_{C_3}$ событий $C_1=\{1, 4\}$, $C_2=\{2, 4\}$, $C_3=\{3, 4\}$. В примере 2 пункта 1 было показано, что случайные переменные f_1, f_2, f_3 зависимы. Их совместное элементарное распределение q не равно произведению элементарных распределений q_1, q_2, q_3 каждой из них: $q\neq q_1\times q_2\times q_3$.

Случайные переменные f_1, f_2, f_3 имеют одинаковые элементарные распределения $q_1=q_2=q_3$ с таблицей значений

x_j	1	0
$q_j(x_j)$	1/2	1/2

Таблица значений для совместного элементарного распределения q случайных переменных f_1, f_2, f_3, как нетрудно проверить, имеет вид:

x	111	110	101	011	100	010	001	000
$q(x)$	1/4	0	0	0	1/4	1/4	1/4	0

Например:

$$q(111) = P(C_1C_2C_3) = P(\{4\}) = 1/4, \quad q(110) = P(C_1C_2C_3') = P(0) = 0,$$

$$q(100) = P(C_1C_2'C_3') = P(\{1\}) = 1/4,$$

$$q(000) = P(C_1'C_2'C_3') = P(0) = 0.$$

Если вместо вероятностных моделей рассматривать механические системы, то сказанное можно пояснить рис. 26.

Зависимость случайных переменных f_1, f_2, f_3 выражается в том, что масса $q(x)$ некоторых точек $x=(x_1, x_2, x_3)$ в пространстве *не равна произведению масс* $q_1(x_1)=q_2(x_2)=q_3(x_3)=1/2$ точек x_1, x_2, x_3 на осях 1, 2, 3.

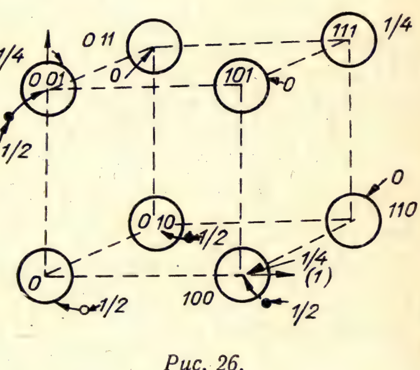


Рис. 26.

§ 4*. ИНДИКАТОРЫ

Среди случайных переменных можно выделить переменные, принимающие только значения 0 и 1. Каждая такая переменная определяется событием, составленным из исходов, для которых она принимает значение 1, и называется индикатором этого события. Разные события имеют разные индикаторы. Поэтому вместо событий можно рассматривать их индикаторы. С другой стороны, в рассматриваемой конечной модели каждая случайная переменная равна некоторой линейной комбинации индикаторов некоторых событий. Часто формальное представление событий и случайных переменных с помощью индикаторов бывает удобно.

Рассмотрим конечную вероятностную модель $(U, \mathscr{A}; p, P)$.

1.4. Определения

Обозначим буквой u произвольный исход из U, а буквой A — произвольное событие из $\mathscr{A}$.

Определение 6. *Индикатором исхода u называется случайная переменная i_u, принимающая значение $i_u(v)=1$ для исхода*

$v=u$ и значение $i_u(v)=0$ для каждого исхода $v\neq u$ из U:

$$i_u(v)=\begin{cases}1 & (v=u),\\ 0 & (v\neq u).\end{cases}$$

Пример 1. Для модели Бернулли $n=1$ испытания таблицы значений индикаторов i_1 и i_0 исходов $u=1$ и $u=0$ имеют вид:

v	1	0
$i_1(v)$	1	0

;

v	1	0
$i_0(v)$	0	1

.

Пример 2. Для модели Бернулли $n=2$ испытаний таблицы значений индикаторов i_{11}, i_{10}, i_{01}, i_{00} исходов 11, 10, 01, 00 приведены в пункте 1 из § 2. Там же рассматривались также другие примеры индикаторов исходов.

Определение 7. *Индикатором события A называется случайная переменная*

$$i_A(v)=\sum_{u\in A} i_u,$$

равная сумме индикаторов i_u исходов u, составляющих событие A.

Ясно, что

$$i_A(v)=\begin{cases}1 & (v\in A),\\ 0 & (v\notin A).\end{cases}$$

Поэтому индикатор i_A события A можно определить как случайную переменную, принимающую значение $i_A(v)=1$ для каждого исхода v из A и значение $i_A(v)=0$ для каждого исхода v не из A.

Пример 1. В модели Бернулли $n=1$ испытания таблицы значений индикаторов i_U, $i_У$, i_H, i_O событий U, $У$, H, O имеют вид:

v	1	0
$i_U(v)$	1	0

;

v	1	0
$i_У(v)$	1	0

;

v	1	0
$i_H(v)$	0	1

;

v	1	0
$i_O(v)$	0	0

.

Пример 2. В модели Бернулли $n=2$ испытаний таблицы значений индикаторов $i_{У_1}$, i_{H_1}, $i_{У_2}$, i_{H_2} событий $У_1$, H_1, $У_2$, H_2 приведены в пункте 1 из § 2. Там же рассматривались также другие примеры индикаторов событий.

2.4. Изоморфизм событий и индикаторов

Обозначим буквой I множество всех индикаторов событий из $\mathscr{A}$ и рассмотрим отображение

$$i:\mathscr{A}\to I$$

класса $\mathscr{A}$ на множество I, которое каждое событие $A \in \mathscr{A}$ отображает в его индикатор $i_A \in I$.

Обозначим буквами A и B произвольные события из $\mathscr{A}$.

Предложение 1. *Отображение i является изоморфизмом класса событий $\mathscr{A}$ на множество индикаторов I и*

(1) $$i_{A+B} = i_A + i_B \quad (AB = 0),$$

(2) $$i_{AB} = i_A \cdot i_B,$$

(3) $$i_{U-A} = i_U - i_A.$$

Доказательство. Если $A \neq B$, то $A - AB \neq 0$ или $B - AB \neq 0$ $i_A(v) = 1 \neq 0 = i_B(v)$ $(v \in A - AB)$ или $i_B(v) = 1 \neq 0 = i_A(v)$ $(v \in B - AB)$ и, значит, $i_A \neq i_B$. Следовательно, отображение $i: \mathscr{A} \to I$ является изоморфизмом.

Равенство (1) вытекает из ассоциативности суммы: если события A и B не пересекаются, то

$$i_{A+B} = \sum_{u \in A+B} i_u = \sum_{u \in A} i_u + \sum_{u \in B} i_u = i_A + i_B.$$

Равенство (2) доказывается равенством значений

$$i_{AB}(v) = i_A(v)\, i_B(v) = \begin{cases} 1 & (v \in AB), \\ 0 & (v \notin AB). \end{cases}$$

Равенство (3) следует из равенства (1):

$$i_U = i_{A+(U-A)} = i_A + i_{U-A}.$$

Предложение 1 доказано.

Замечание. Если постоянную i_U со значением 1 обознач символом 1, то равенство (3) можно записать также в форме

(3′) $$i_{A'} = 1 - i_A$$

3.4. Стандартное представление случайной переменной

Рассмотрим произвольную случайную переменную f, определенную на множестве исходов U.

Из определения индикатора i_u произвольного исхода u из множества U следует, что

$$f(v) = \sum_{u \in U} f(u)\, i_u(v)$$

для каждого исхода v из множества U:

$$f(u)\, i_u(v) = \begin{cases} f(v) & (u = v), \\ 0 & (u \neq v). \end{cases}$$

Это значит, что случайная переменная f равна линейной комбинации индикаторов i_u с коэффициентами $f(u)$:

(1) $$f = \sum_{u \in U} f(u)\, i_u.$$

Примеры. См. примеры 1—3 пункта 1 из § 2.

Обозначим множество значений случайной переменной f буквой X:

$$X = f(U).$$

Объединяя в сумме (1) слагаемые с одинаковыми коэффициентами $f(u) = x$ для каждого значения x переменной f, получаем представление f в виде линейной комбинации индикаторов событий:

$$f^{-1}(x) = \{u : f(u) = x\}.$$

Предложение 2. *Для каждой случайной переменной f с множеством значений X верно равенство*

$$f = \sum_{x \in X} x \cdot i_{f^{-1}(x)} \tag{2}$$

Доказательство. В самом деле,

$$f = \sum_{u \in U} f(u)\, i_u = \sum_{x \in X} \left(\sum_{u \in f^{-1}(x)} x \cdot i_u \right) =$$

$$= \sum_{x \in X} \left(x \sum_{u \in f^{-1}(x)} i_u \right) = \sum_{x \in X} x \cdot i_{f^{-1}(x)}$$

Условимся называть равенство (2) стандартным представлением случайной переменной f в виде линейной комбинации индикаторов или, коротко, стандартным представлением для f.

Примеры. См. примеры 1—3 пункта 1 из § 2.

4.4. Специальный пример

Рассмотрим произвольные случайные переменные f_1 и f_2 на множестве исходов U. Как и в § 3, обозначим буквами X_1 и X_2, q_1 и q_2 множества значений и элементарные распределения случайных переменных f_1 и f_2. Рассмотрим также пару $f = (f_1, f_2)$, ее множество значений $X = X_1 \times X_2$ и совместное элементарное распределение q случайных переменных f_1 и f_2.

В § 3 было доказано, что независимость случайных переменных f_1 и f_2 эквивалентна равенству

$$q(f(u)) = q_1(f_1(u))\, q_2(f_2(u)) \tag{*}$$

для каждого исхода $u \in U$.

Предположим дополнительно, что

$$q(x) = q(x_1, x_2) > 0$$

для каждого значения $x = (x_1, x_2)$ случайной переменной $f = (f_1, f_2)$. В этом случае

$$q_1(x_1) = \sum_{x_1 \in X_1} q(x_1, x_2) > 0,$$

$$q_2(x_2) = \sum_{x_2 \in X_2} q(x_1, x_2) > 0$$

для каждых значений x_1 и x_2 случайных переменных f_1 и f_2. И, следовательно, равенство (*) эквивалентно равенству

$$(**)\qquad \log \frac{q(f(u))}{q_1(f_1(u))\, q_2(f_2(u))} = 0$$

для каждого исхода u (логарифм берется по основанию строго большему 1).

Таким образом, при сделанных предположениях о строгой положительности вероятностей значений независимость случайных переменных f_1 и f_2 эквивалентна равенству (**).

Рассмотрим случайную переменную

$$i = \sum_{u \in U} \log \frac{q(f(u))}{q_1(f_1(u))\, q_2(f_2(u))}\, i_u.$$

Она описывает зависимость случайных переменных f_1 и f_2: они независимы, если и только если i тождественно равна 0.

Рассмотрим произвольные значения $x \in X_1$, $x_2 \in X_2$ и $x = (x_1, x_2) \in X$. Так как

$$i(u) = \log \frac{q(x)}{q_1(x_1)\, q_2(x_2)}$$

для каждого $u \in f^{-1}(x)$, то

$$i = \sum_{x \in X} \log \frac{q(x)}{q_1(x_1)\, q_2(x_2)}\, i_{f^{-1}(x)}\ .$$

Замечание. Это равенство является стандартным представлением для i, только если отношения $q(x)/q_1(x_1)q_2(x_2)$ различны для различных значений $x = (x_1, x_2)$.

Пример 1. Рассмотрим модель Бернулли $n=2$ испытаний с вероятностью успеха $a(0<a<1)$ и случайные переменные $f_1 = i_{y_1}$, $f_2 = i_{y_2}$. Имеем:

$$i = \log \frac{q(1,1)}{q_1(1)\, q_2(1)} i_{11} + \log \frac{q(1,0)}{q_1(1)\, q_2(0)} i_{10} + \log \frac{q(0,1)}{q_1(0)\, q_2(1)}\, i_{01} + \\ + \log \frac{q(0,0)}{q_1(0)\, q_2(0)}\, i_{00} = \log \frac{a^2}{a.a}\, i_{11} + \log \frac{a(1-a)}{a(1-a)} i_{10} + \\ + \log \frac{(1-a)\,a}{(1-a)\,a} i_{01} + \log \frac{(1-a)^2}{(1-a)(1-a)}\, i_{00} = 0.$$

Пример 2. Рассмотрим модель Бернулли $n=2$ испытаний с вероятностью успеха $a(0<a<1)$ и случайные переменные f_1, f_2 со значениями

$$f_1(u) = u_1,\ f_2(u) = u_1 + u_2$$

для каждого исхода $u = (u_1, u_2)$. Имеем:

$$i = \log \frac{q(1,2)}{q_1(1)\, q_2(2)}\, i_{11} + \log \frac{q(1,1)}{q_1(1)\, q_2(1)}\, i_{10} + \log \frac{q(0,1)}{q_1(0)\, q_2(1)}\, i_{01} + \\ + \log \frac{q(0,0)}{q_1(0)\, q_2(0)}\, i_{00} = \log \frac{a^2}{a \cdot a^2} i_{11} + \log \frac{a(1-a)}{a \cdot 2a(1-a)}\, i_{10} +$$

$$+\log\frac{(1-a)a}{(1-a)2a(1-a)}i_{01}+\log\frac{(1-a)^2}{(1-a)(1-a)^2}i_{00}=$$

$$=\log\frac{1}{a}i_{11}+\log\frac{1}{2a}i_{10}+\log\frac{1}{2(1-a)}i_{01}+\log\frac{1}{1-a}i_{00}.$$

В частности, если $a=1/2$, то

$$i=\log 2(i_{11}+i_{00}).$$

Если основание логарифма равно 2, то

$$i=i_{11}+i_{00}.$$

Пример 3. Случайная переменная i для индикаторов $f_1=i_A$ и $f_2=i_B$ событий A и B, для которых $P(AB)$, $P(AB')$, $P(A'B)$, $P(A'B')>0$, выражается равенством:

$$i=\log\frac{P(AB)}{P(A)P(B)}i_{AB}+\log\frac{P(AB')}{P(A)P(B')}i_{AB'}+$$

$$+\log\frac{P(A'B)}{P(A')P(B)}i_{A'B}+\log\frac{P(A'B')}{P(A')P(B')}i_{A'B'}.$$

Если события A и B независимы, то события A и B', A' и B, A' и B' *независимы* и $i=0$. Если $i\neq 0$, то все эти события *зависимы*.

Рассматривавшийся специальный пример будет использован в следующей главе при описании информации случайных переменных.

Глава 2

СРЕДНЕЕ СЛУЧАЙНОЙ ПЕРЕМЕННОЙ

Важнейшей числовой характеристикой случайной переменной является ее *среднее вероятностное*, определяемое как сумма произведений вероятностей исходов на соответствующие значения случайной переменной. В механической модели среднее вероятностное интерпретируется как *центр масс* рассматриваемой системы точек. Если все исходы равновероятны, то среднее вероятностное равно среднему арифметическому.

С помощью среднего вероятностного определяются другие важные числовые характеристики случайных переменных: *дисперсия*, *коэффициент корреляции*, *информация* и *энтропия*.

При первом чтении рекомендуется ограничиться параграфами 1—3.

Задачи к главе 2 собраны в § 1 главы 2 части III.

§ 1. ПРИМЕРЫ

Рассмотрим несколько примеров, разъясняющих понятие среднего вероятностного.

1.1. Пример 1. Игра с костью

Игральная кость подбрасывается один раз. Если выпадает четное число очков, то игрок выигрывает с рублей, а если нечетное — проигрывает. Каково среднее вероятностное выигрыша игрока?

1.1.1. Первая модель

Как было объяснено в примере 1 § 1 главы 1, подбрасывание кости один раз описывается стандартной моделью Лапласа $n=6$ равновероятных исходов 1, 2, 3, 4, 5, 6, а выигрыш игрока — случайной переменной f на множестве исходов $U=\{1, 2, 3, 4, 5, 6\}$, значение $f(u)$ которой для каждого исхода u определяется таблицей

u	1	2	3	4	5	6
$f_1(u)$	$-c$	$+c$	$-c$	$+c$	$-c$	$+c$

Если выпадает u очков, то игрок выигрывает $f(u)$ рублей.

Элементарная вероятность $p(u)$ каждого исхода u равна 1/6.

Среднее вероятностное $E(f)$ выигрыша f выражается равенствами:

$$E(f)=\Sigma f(u)p(u)=(-c)1/6+(+c)1/6+(-c)1/6+(+c)1/6+$$
$$+(-c)1/6+(+c)1/6=(-c+c-c+c-c+c)1/6=0.$$

В данном случае среднее вероятностное равно *среднему арифметическому* значений случайной переменной f.

Среднее вероятностное $E(f)$ оценивает *ожидаемый выигрыш*. В самом деле, вероятность $p(u)$ исхода u оценивает реализуемость исхода u, его долю в достоверном событии U. Поэтому естественно рассчитывать на долю $p(u)$ выигрыша $f(u)$, получаемого при реализации исхода u. Следовательно, можно для каждого исхода u оценить ожидаемый выигрыш произведением $f(u)p(u)$. Общий ожидаемый выигрыш будет оцениваться тогда суммой всех таких произведений, т. е. средним вероятностным $E(f)$ случайной переменной f.

Равенство ожидаемого выигрыша нулю можно истолковать как *безобидность* рассматриваемой игры.

2.1.1. Вторая модель

Для решения задачи можно использовать также модель с множеством исходов $X=\{-c, +c\}$ и элементарной вероятно-

стью q со значениями $q(-c)=q(+c)=1/2$. Игрок с вероятностью $q(-c)=1/2$ выигрывает $-c$ рублей, а с вероятностью $q(+c)=$ $=1/2$ — $+c$ рублей.

Среднее вероятностное выигрыша выражается равенствами:

$$\Sigma xq(x)=-c\cdot 1/2+ +c\cdot 1/2=0.$$

Равенство

$$E(f)=\Sigma xq(x)$$

объясняется тем, что элементарная вероятность q является элементарным распределением случайной переменной f и

$$q(-c)=P(\{u : f(u)=-c\})=p(1)+p(3)+p(5)=1/2,$$
$$q(+c)=P(\{u:f(u)=+c\})=p(2)+p(4)+p(6)=1/2.$$

Поэтому

$$E(f)=\Sigma f(u)p(u)=f(1)p(1)+f(2)p(2)+f(3)p(3)+f(4)p(4)+$$
$$+f(5)p(5)+f(6)p(6)=-cq(-c)+(+c)q(+c)=\Sigma xq(x).$$

2.1. Пример 2. Игра с монетой

Симметричная монета подбрасывается два раза. Если при первом подбрасывании выпадает герб, то игрок выигрывает 5 копеек, а если цифра — проигрывает. То же при втором подбрасывании. Каково среднее вероятностное выигрыша игрока?

1.2.1. Первая модель

Как было объяснено в примере 2 из § 1 главы 1, подбрасывание симметричной модели два раза описывается моделью Бернулли $n=2$ испытаний с вероятностью успеха $a=1/2$, а выигрыш игрока при первом и втором подбрасывании и общий — случайными переменными $f_1(u)$, $f_2(u)$ и $f(u)$ на множестве исходов $U=$ $=\{11, 10, 01, 00\}$, значения $f_1(u)$, $f_2(u)$ и $f(u)$ которых для каждого исхода u определяются соответственно таблицами:

u	11	10	01	00
$f_1(u)$	+5	+5	−5	−5

;

u	11	10	01	00
$f_2(u)$	+5	−5	+5	−5

;

u	11	10	01	00
$f(u)$	+10	0	0	−10

.

Элементарная вероятность $p(u)$ каждого исхода u равна 1/4.

Средние вероятностные $E(f_1)$, $E(f_2)$ и $E(f)$ выигрышей f_1, f_2 и f выражаются равенствами:

$$E(f_1)=\Sigma f_1(u)p(u)=f_1(11)p(11)+f_1(10)p(10)+f_1(01)p(01)+$$
$$+f_1(00)p(00)=(+5)(1/4)+(+5)(1/4)+(-5)(1/4)+(-5)(1/4)=$$
$$=(+5+5-5-5)(1/4)=0,$$
$$E(f_2)=\Sigma f_2(u)p(u)=(+5-5+5-5)(1/4)=0,$$
$$E(f)=\Sigma f(u)p(u)=(+10+0+0-10)(1/4)=0.$$

Как и в примере 1, средние вероятностные равны средним арифметическим.

Заметим, что

$$f=f_1+f_2,\ E(f)=E(f_1)+E(f_2).$$

Это равенство не случайно:

$$\Sigma f(u)p(u)=f(11)p(11)+f(10)p(10)+f(01)p(01)+f(00)p(00)=$$
$$=(f_1(11)+f_2(11))p(11)+(f_1(10)+f_2(10))p(10)+$$
$$+(f_1(01)+f_2(01))p(01)+(f_1(00)+f_2(00))p(00)=$$
$$=\Sigma f_1(u)p(u)+\Sigma f_2(u)p(u).$$

2.2.1. Вторая модель

Для решения задачи можно использовать также модель с множеством исходов $X=\{-10,\ 0,\ +10\}$ и элементарной вероятностью q, значение которой для каждого исхода x определяется таблицей

x	-10	0	$+10$
$q(x)$	$1/4$	$1/2$	$1/4$

Игрок с вероятностью $q(-10)=1/4$ выигрывает -10 копеек, с вероятностью $q(10)=1/2$ — 0 копеек и с вероятностью $q(+10)$ — $+10$ копеек.

Среднее вероятностное выигрыша выражается равенствами:

$$\Sigma xq(x)=(-10)(1/4)+0(1/2)+(+10)(1/4)=0.$$

Равенство

$$E(f)=\Sigma xq(x)$$

следует из того, что элементарная вероятность q является элементарным распределением случайной переменной f.

3.1. Пример 3. Выстрел по мишени

Производится один выстрел по стандартной мишени. Известны вероятности получения каждого числа очков. Каково среднее вероятностное числа очков?

Как было объяснено в примере 3 из § 1 главы 1, один выстрел по стандартной мишени описывается моделью с множеством исходов $U=\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ и элементарной вероятностью p, которая предполагается известной, а получаемое число очков — тождественной случайной переменной f. Рассмотрим ту же элементарную вероятность p, что и в примере 1 из § 1 главы 1, значение $p(u)$ которой для каждого исхода определяется таблицей:

u	0	1	2	3	4	5	6	7	8	9	10
$p(u)$	0,01	0,01	0,01	0,01	0,01	0,05	0,05	0,05	0,10	0,20	0,50

Среднее вероятностное $E(f)$ числа очков f выражается равенствами:

$$E(f)=\Sigma f(u)p(u)=0\cdot p(0)+1\cdot p(1)+2\cdot p(2)+\ldots+10p(10)=$$
$$=0\cdot 0{,}01+1\cdot 0{,}01+2\cdot 0{,}01+3\cdot 0{,}01+4\cdot 0{,}01+5\cdot 0{,}05+6\cdot 0{,}05+7\cdot 0{,}05+$$
$$+8\cdot 0{,}10+9\cdot 0{,}20+10\cdot 0{,}50=8{,}60.$$

В данном случае среднее вероятностное *не равно* среднему арифметическому значений случайной переменной.

Среднее вероятностное $E(f)=8{,}60$ оценивает ожидаемое число очков. Заметим, что это число не является значением случайной переменной f.

§ 2. ОПРЕДЕЛЕНИЯ

Среднее вероятностное случайной переменной определяется как сумма произведений ее значений на вероятности соответствующих исходов. Оно оценивает ожидаемое значение случайной переменной и называется также ее *математическим ожиданием*.

Основными свойствами среднего вероятностного являются его положительность, нормированность и линейность. Эти свойства аналогичны основным свойствам вероятности.

1.2. Определение среднего

Рассмотрим конечную вероятностную модель $(U, \mathscr{A}; p, P)$, случайную переменную f на U, ее множество значений X и элементарное распределение q.

Определение 1. *Средним вероятностным случайной переменной называется число*

$$E(f)=\sum_{u\in U} f(u)\,p(u).$$

Примеры 1—3 из § 1 поясняют это определение. Условимся также писать

$$E(f)=\Sigma f(u)\,p(u).$$

Пример. Для модели Лапласа с $n=n(U)$ равновероятными исходами среднее вероятностное $E(f)$ для каждой случайной переменной f равно ее среднему арифметическому:

$$E(f)=\frac{1}{n(U)}\sum_{u\in U} f(u).$$

Обозначим множество всех случайных переменных на U буквой $\mathcal{F}$.

Функция E на множестве $\mathcal{F}$, значением которой для каждой случайной переменной f из $\mathcal{F}$ является ее среднее вероятностное $E(f)$, называется средним вероятностным.

Из определения элементарного распределения q следует, что среднее вероятностное случайной переменной f равно сумме произведений ее значений x на их вероятности $q(x)$:

Предложение 1. $E(f)=\sum\limits_{x\in X} xq(x)$.

Доказательство. По определению,

$$f^{-1}(x)=\{u: f(u)=x\},\quad q(x)=P\left(f^{-1}(x)\right)=\sum_{u\in f^{-1}(x)} p(u)$$

для каждого значения $x\in X$. Используя эти равенства и группируя слагаемые в сумме, определяющей $E(f)$, получаем:

$$E(f)=\sum_{u\in U} f(u)\,p(u)=\sum_{x\in X}\Bigl(\sum_{u\in f^{-1}(x)} f(u)\,p(u)\Bigr)=$$

$$=\sum_{x\in X}\Bigl(\sum_{u\in f^{-1}(x)} xp(u)\Bigr)=\sum_{x\in X}\Bigl(x\cdot\sum_{u\in f^{-1}(x)} p(u)\Bigr)=$$

$$=\sum_{x\in X} xp\left(f^{-1}(x)\right)=\sum_{x\in X} xq(x).$$

Предложение 1 доказано.

Условимся также писать

$$E(f)=\Sigma xq(x).$$

Замечание. Доказанное равенство часто используется в качестве определения среднего вероятностного.

Из предложения 1 вытекает несколько следствий.

Первое из этих следствий показывает, что среднее вероятностное определяется *распределением* случайной переменной.

Следствие. *Если случайные переменные имеют одинаковые распределения, то их средние вероятностные равны.*

Доказательство. Из равенства распределений случайных пепеременных следуют равенства их множеств значений и элементарных распределений. А из предложения 1 вытекает, что если случайные переменные f_1 и f_2 на множестве исходов U имеют одинаковые множества значений $X_1 = X_2$ и элементарные распределения $q_1 = q_2$, то

$$E(f_1) = \Sigma x_1 q_1(x_1) = \Sigma x_2 q_2(x_2) = E(f_2).$$

Второе следствие утверждает, что среднее вероятностное *постоянной* равно ее значению. Рассмотрим произвольное число c. Постоянную на множестве исходов U со значением c условимся обозначать буквой c.

Правило постоянной:

$$E(c) = c.$$

Доказательство. Если $f(u) = c$ для каждого исхода u, то $X = \{c\}$, $q(c) = 1$ и $E(f) = c \cdot 1 = c$.

Замечание. Если $f = c$ почти всюду, т. е. $q(x) = 0$ для каждого значения $x \neq c$ случайной переменной f, то

$$E(f) = cq(c) + \sum_{x \neq c} xq(x) = c \cdot 1 + 0 = c.$$

Третье следствие утверждает, что среднее вероятностное *индикатора* события равно вероятности этого события. Рассмотрим произвольное событие $A \subseteq U$ и его индикатор i_A.

Правило индикатора:

$$E(i_A) = P(A).$$

Доказательство. Если $f = i_A$, то $X = \{1, 0\}$, $q(1) = P(A)$, $q(0) = P(A')$, $E(f) = 1 \cdot q(1) + 0 \cdot q(0) = P(A)$.

Условимся в дальнейшем среднее вероятностное коротко называть также *средним* случайной переменной.

Замечание. Правило индикатора показывает, что аналогично тому, как события можно отождествить с их индикаторами, вероятности событий можно отождествить со средними их индикаторов. В этом смысле понятия случайной переменной и ее среднего обобщают соответственно понятия события и его вероятности.

Пример. По прямоугольной решетке, составленной из сантиметровых квадратов, ползает муха. В каждом узле решетки она поворачивает влево или вправо, выбирая направление наугад. Возможные пути мухи из угла A в угол B по четырем отрезкам решетки, ориентированным относительно первого пройденного ею угла, изображены на рис. 27.

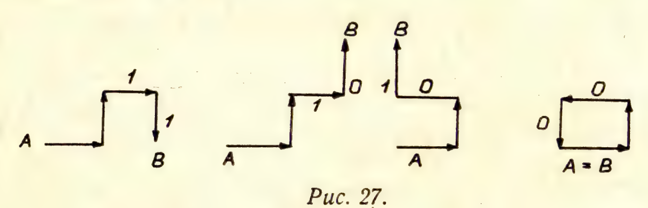


Рис. 27.

Каково среднее расстояние между A и B?

Движение мухи можно описать моделью Бернулли $n=2$ испытаний с вероятностью успеха $a=1/2$: соответствие исходов и возможных путей указано на рисунке (1 описывает поворот вправо, 0 — влево).

Расстояние между узлами A и B описывает случайная переменная d с таблицей значений

u	11	10	01	00
$d(u)$	2	$2\sqrt{2}$	2	0

Стандартное представление случайной переменной d имеет вид:

$$d=2\sqrt{2}\cdot i_{\{10\}}+2\cdot i_{\{11,01\}}.$$

Среднее для d выражается равенствами:

$$E(d)=2\cdot 1/4+2\sqrt{2}\cdot 1/4+2\cdot 1/4+0\cdot 1/4=2\sqrt{2}\cdot 1/4+2\cdot 1/2=$$

$$=2\sqrt{2}\cdot P(\{10\})+2\cdot P(\{11,\ 01\})=2\sqrt{2}\cdot E(i_{\{10\}})+2\cdot E(i_{\{11,01\}})=$$

$$=1+\sqrt{2}/2.$$

Замечание. Равенства

$$P(\{u:d(u)=0)=P(\{00\})=1/4$$

выражают вероятность того, что муха вернется в исходную точку.

2.2. Содержание и механическая интерпретация

Среднее $E(f)$ случайной переменной f с содержательной точки зрения можно считать *оценкой ожидаемого значения* случайной переменной f. Это можно обосновать следующим образом. Вероятность $p(u)$ исхода u оценивает реализуемость исхода u, его *долю* в достоверном событии U. Поэтому можно считать, что произведение $f(u)\cdot p(u)$ оценивает *долю* значения $f(u)$ для исхода u в ожидаемом значении случайной переменной f. Само это значение естественно оценивать тогда суммой всех таких произведений, т. е. средним $E(f)$.

Примеры 1—3 из § 1 поясняют это рассуждение.

Из предложения 1 следует, что в механической системе точек x с массами $q(x)$ среднее $E(f)$ является *центром масс.*

Пример 1. В примере 1 из § 1 в симметричных относительно нуля точках $x=-c$ и $x=+c$ расположены одинаковые массы $q(-c)=1/2$ и $q(+c)=1/2$. Центром масс является точка 0 (рис. 28).

Рис. 28.

Пример 2. В примере 2 из § 1 в точке $x=0$ расположена масса $q(0)=1/2$, а в симметричных относительно нуля точках $x=-10$ и $x=+10$ расположены одинаковые массы $q(-10)=1/4$ и $q(+10)=1/4$. Центром масс является точка 0 (рис. 29).

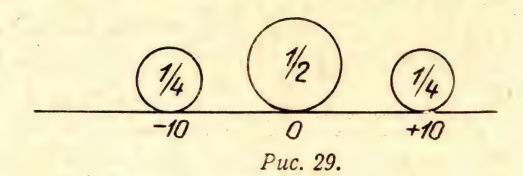


Рис. 29.

Пример 3. В примере 3 из § 1 расположение масс $q(x)=p(u)$ в точках $x=u$ определяется соответствующей таблицей. Центром масс является точка 8, 6 (рис. 30).

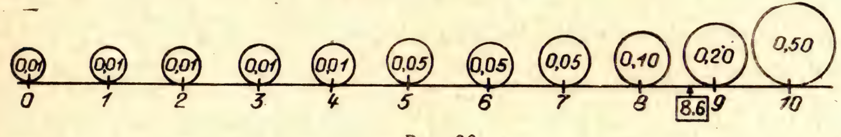


Рис. 30.

3.2 Основные свойства среднего

Основными свойствами среднего являются его *положительность*, *нормированность* и *линейность*. Последнее из этих свойств выражается правилами сложения и умножения на число. Основные свойства среднего обобщают основные свойства вероятности.

Условимся говорить, что функция E на множестве $\mathcal{F}$ *положительна*, если и только если ее значение $E(f)$ положительно для каждой случайной переменной f, все значения $f(u)$ которой положительны. Т. е. положительность E означает, что $E(f) \geqslant 0$ для каждой случайной переменной $f \geqslant 0$.

1-е основное свойство. *Среднее E положительно.*

Доказательство. Положительность среднего следует из положительности вероятности. Так как $p(u) \geqslant 0$ для каждого исхода u, то для каждой случайной переменной f, для которой $f(u) \geqslant 0$ для каждого исхода u,

$$E(f)=\Sigma f(u)\,p(u) \geqslant 0.$$

Замечание. Положительность функции E *не означает*, что все ее значения $E(f)$ положительны.

Условимся говорить, что функция E на $\mathcal{F}$ *нормирована*, если и только если ее значение $E(1)$ для постоянной 1 равно числу 1:

$$E(1)=1.$$

2-е основное свойство. *Среднее E нормировано.*

Доказательство. Нормированность среднего следует из нормированности вероятности:

$$E(1)=\Sigma p(u)=1.$$

Замечание. Нормированность среднего вытекает также из следствия 2 предложения 1.

Условимся говорить, что для функции E на $\mathcal{F}$ верно *правило сложения*, если значение $E(f+g)$ функции E для суммы $f+g$ каждых случайных переменных f и g на U равно сумме $E(f)+$ $+E(g)$ значений $E(f)$ и $E(g)$ функции E для случайных переменных f и g.

Правило сложения:

$$E(f+g)=E(f)+E(g).$$

Условимся говорить, что для функции E на $\mathcal{F}$ верно *правило умножения на число*, если значение $E(c\cdot f)$ для произведения $c\cdot f$ каждой случайной переменной f на U и каждого числа c равно произведению $c\cdot E(f)$ значения $E(f)$ функции E для случайной переменной f и числа c.

Правило умножения на число:

$$E(c\cdot f)=c\cdot E(f).$$

Если для функции E на $\mathcal{F}$ верно правило сложения, то говорят, что она *аддитивна*. Если для функции E на $\mathcal{F}$ верно правило умножения на число, то говорят, что она *однородна*. Если функция E на $\mathcal{F}$ аддитивна и однородна, то говорят, что она *линейна*.

Определение. *Функция E на множестве $\mathcal{F}$ случайных переменных называется линейной, если и только если для нее верны правила сложения и умножения на число.*

Это определение позволяет сформулировать

3-е основное свойство. *Среднее E линейно.*

Доказательство. Линейность среднего E следует из линейности суммы. Для каждых случайных переменных f и g на U и числа c верны равенства:

$$1)\qquad E(f+g)=\Sigma(f(u))+g(u))p(u)=\Sigma f(u)p(u)+\Sigma g(u)p(u)=$$

$$=E(f)+E(g),$$

$$2)\qquad E(c\cdot f)=\Sigma(c\cdot f(u)p(u)=c\cdot\Sigma f(u)p(u)=c\cdot E(f).$$

Значит, среднее E линейно.

Три основные свойства среднего выражает объединяющая эти свойства

Теорема. *Среднее E является положительной, нормированной и линейной функцией на множестве случайных переменных $\mathcal{F}$.*

Замечание. Из правила индикатора и теоремы об основных свойствах среднего вытекает теорема об основных свойствах вероятности.

Все общие свойства среднего являются следствиями его положительности, нормированности и линейности. Предложения, выражающие эти общие свойства, являются следствиями доказанной теоремы.

4.2. Следствия основных свойств

Из основных свойств среднего вытекают, в частности, *правило вычитания*, *правило неравенства* и *общее правило линейности*.

Рассмотрим произвольные случайные переменные f и g на U.

Правило вычитания:

$$E(g-f)=E(g)-E(f).$$

Доказательство. Из линейности среднего E следует, что

$$E(g-f)=E(g+(-f))=E(g)+E(-f)=E(g)-E(f).$$

Условимся говорить, что случайная переменная f меньше случайной переменной g, и писать $f\leqslant g$, если и только если $f(u)\leqslant g(u)$ для каждого исхода u.

Правило неравенства:

$$E(f)\leqslant E(g) \qquad (f\leqslant g).$$

Доказательство. Из правила вычитания и положительности среднего E следует, что

$$E(g)-E(f)=E(g-f)\geqslant 0,$$

если $f\leqslant g$ и $g-f\geqslant 0$.

Рассмотрим произвольные семейство (f_j) случайных переменных f_j на U и семейство (c_j) чисел c_j, имеющие одно и то же конечное множество индексов I.

Общее правило линейности:

$$E(\Sigma c_j f_j)=\Sigma c_j E(f_j).$$

Доказательство. Рассмотрим множество M всех номеров m таких, что общее правило сложения верно для каждых семейства $(f_j)_{j\in I}$ случайных переменных f_j на U и семейства $(c_j)_{j\in I}$ чисел c_j, множество индексов I которых составлено из $m+1$ элементов.

1. Номер $0\in M$;

2. Для каждого номера n верно предложение: если $n\in M$, то $n+1\in M$. Действительно, рассмотрим произвольные семейство $(f_j)_{j\in\bar{I}}$ случайных переменных f_j на U и семейство $(c_j)_{j\in\bar{I}}$ чисел c_j, множество индексов I которых составлено из $(n+1)+1$ элементов. Для каждого индекса $j_0\in\bar{I}$ определены семейства $(f_j)_{j\in I}$ и $(c_j)_{j\in I}$ с множеством индексов $I=\bar{I}-\{i_0\}$, составленным из $n+1$ элементов. Положим:

$$f=c_{j_0}f_{j_0}, \quad g=\sum_{j\in I}c_j f_j.$$

Вследствие линейности среднего E имеем:

$$E=\left(\sum_{j\in\bar{I}}c_j f_j\right)=E(f+g)=E(f)+E(g)=$$

$$=c_{j_0}E(f_{j_0})+E\left(\sum_{j\in I}c_j f_j\right).$$

Если $n \in M$, то

$$E\left(\sum_{j\in I} c_j f_j\right) = \sum_{j\in I} c_j E(f_j)$$

и, следовательно,

$$E\left(\sum_{j\in \bar{I}} c_j f_j\right) = c_{j_0} E(f_{j_0}) + \sum_{j\in I} c_j E(f_j) = \sum_{j\in \bar{I}} c_j E(f_j),$$

т. е. $n+1 \in M$.

По принципу индукции из пунктов 1 и 2 следует, что множество M равно множеству N всех натуральных чисел. Общее правило линейности доказано.

Из общего правила линейности следует

Общее правило сложения:

$$E(\Sigma f_j) = \Sigma E(f_j).$$

Замечание. В свою очередь, общее правило линейности следует из общего правила сложения и правила умножения на число для среднего E.

Из правил индикатора, вычитания, неравенства и общего правила сложения для среднего E вытекают *правило вычитания*, *правило неравенства* и *общее правило сложения* для вероятности P.

5.2. Примеры

Рассмотрим несколько примеров, поясняющих использование правил для среднего E.

Пример 1. Рассмотрим модель Бернулли n испытаний с вероятностью успеха a и случайную переменную s, равную сумме индикаторов $s_j = i_{У_j}$ событий $У_j$, описывающих успех при j-м испытании:

$$s = \sum_{1\leqslant j\leqslant n} s_j.$$

Случайная переменная s описывает *общее число успехов* при n испытаниях.

Используя общее правило сложения для среднего, правило индикатора и одинаковость испытаний в схеме Бернулли, получаем:

$$E(s) = \sum_{1\leqslant j\leqslant n} E(s_j) = \sum_{1\leqslant j\leqslant n} P(У_j) = n\cdot a.$$

Для сравнения получим это же равенство, непосредственно используя предложение 1. Как было отмечено в пункте 6 из § 2 главы 1, случайная переменная s имеет биномиальное распределение:

$$q(m) = \binom{n}{m} a^m (1-a)^{n-m}$$

для каждого номера $m=0, \ldots, n$. Поэтому

$$E(s) = \sum_{0 \leqslant m \leqslant n} mq(m) = \sum_{0 \leqslant m \leqslant n} m \binom{n}{m} a^m (1-a)^{n-m} =$$

$$= \sum_{0 \leqslant m \leqslant n} n \binom{n-1}{m-1} a^m (1-a)^{n-m} =$$

$$= n \cdot a \cdot \sum_{1 \leqslant m \leqslant n} \binom{n-1}{m-1} a^{m-1} (1-a)^{(n-1)-(m-1)} =$$

$$= n \cdot a (a + 1 - a)^{n-1} = n \cdot a.$$

Пример 2. Для модели Бернулли n испытаний с вероятностью успеха a рассмотрим случайную переменную s/n, описывающую *частоту успеха*. Используя результат примера 1 и линейность среднего E, получаем:

$$E(s/n) = E(s)/n = n \cdot a/n = a.$$

В модели Бернулли при любом числе испытаний n среднее частоты успеха равно вероятности успеха a.

Пример 3. Симметричная монета подбрасывается сто раз. Каково ожидаемое число появления герба?

Подбрасывание симметричной монеты сто раз описывается моделью Бернулли $n=100$ испытаний с вероятностью успеха $a=1/2$. Задача сводится к вычислению среднего для числа успехов. Имеем:

$$E(s) = n \cdot a = 100 \cdot 1/2 = 50.$$

Замечание. В то же время вероятность появления ровно 50 гербов при 100 подбрасываниях симметричной монеты меньше, чем 0,08.

6.2. Правило умножения

Среднее произведения *независимых* случайных переменных равно произведению их средних. Существуют зависимые случайные переменные, среднее произведения которых *не равно* произведению их средних.

Рассмотрим независимые случайные переменные f и g на U.

Правило умножения:

$$E(f \cdot g) = E(f) \cdot E(g).$$

Доказательство. Правило умножения для среднего следует из правила умножения для вероятности, правила индикатора и общего правила линейности для среднего.

Для каждых независимых событий A и B из правила умножения для вероятности и правила индикатора для среднего вытекают равенства:

$$E(i_A i_B) = E(i_{AB}) = P(AB) = P(A)P(B) = E(i_A)E(i_B).$$

Обозначим множества значений переменных f и g буквами X и Y. Независимость случайных переменных f и g означает независимость событий $A=f^{-1}(x)$ и $B=g^{-1}(y)$ для каждых $x\in X$ и $y\in Y$. Используя стандартные представления случайных переменных f и g, полученное равенство для индикаторов и общее правило линейности для среднего, убеждаемся в том, что

$$E(f\cdot g)=E\left(\sum_{x\in X} xi_{f^{-1}(x)} \sum_{y\in Y} yi_{g^{-1}(y)}\right)=$$

$$=E\left(\sum_{x\in X,\, y\in Y} xyi_{f^{-1}(x)}\, i_{g^{-1}(y)}\right)=\sum_{x\in X,\, y\in Y} xyE\left(i_{f^{-1}(x)}\, i_{g^{-1}(y)}\right)=$$

$$=\sum_{x\in X,\, y\in Y} xyE\left(i_{f^{-1}(x)}\right)E\left(i_{g^{-1}(y)}\right)=$$

$$=E\left(\sum_{x\in X} x\cdot i_{f^{-1}(x)}\right)=E\left(\sum_{y\in Y} y\cdot i_{g^{-1}(y)}\right)=E(f)\,E(g).$$

Правило умножения для среднего E доказано.

Замечание. Из правила умножения для среднего и правила индикатора, в свою очередь, вытекает правило умножения для вероятности.

Рассмотрим семейство (f_j) независимых случайных переменных f_j на U, имеющее конечное множество индикаторов I.

Общее правило умножения:

$$E(\Pi f_j)=\Pi E(f_j).$$

Доказательство. Рассмотрим множество M всех номеров m таких, что общее правило умножения верно для каждого семейства $(f_j)_{j\in I}$ независимых случайных переменных f_j на U, множество индексов I которого составлено из $m+1$ элементов.

1. Номер $0\in M$;

2. Для каждого номера n верно предложение: если $n\in M$, то $n+1\in M$. Действительно, рассмотрим произвольное семейство $(f_j)_{j\in\bar{I}}$ независимых случайных переменных f_j на U, множество индексов $\bar{I}$ которого составлено из $(n+1)+1$ элементов. Для каждого индекса $j_0\in\bar{I}$ определено семейство $(f_j)_{j\in I}$ независимых случайных переменных f_j на U, множество индексов $I=\bar{I}-\{j_0\}$ которого составлено из $n+1$ элементов. Положим:

$$f=f_{j_0},\quad g=\prod_{j\in I} f_j.$$

Из независимости случайных переменных f_j семейства $(f_j)_{j\in\bar{I}}$ вытекает независимость случайных переменных f и g. Следовательно, по правилу умножения,

$$E(f\cdot g)=E(f)\cdot E(g).$$

Если $n\in M$, то

$$E(g)=\prod_{j\in I} E(f_j)$$

и, значит,

$$E\left(\prod_{j\in\bar{I}} f_j\right) = E(f\cdot g) = E(f)\cdot E(g) = E(f_{j_0})\prod_{j\in I} E(f_j) = \prod_{j\in\bar{I}} E(f_j),$$

т. е. $n+1\in M$.

По принципу индукции из пунктов 1 и 2 следует, что множество M равно множеству N всех натуральных чисел. Общее правило умножения доказано.

В следующих примерах 1—3 рассматривается *модель Бернулли $n=2$ испытаний с вероятностью успеха a.*

Пример 1. Случайные переменные $f=s_1=i_{Y_1}$ и $g=s_2=i_{Y_2}$ *независимы*. Для них

$$E(f\cdot g)=P(Y_1Y_2)=P(Y_1)P(Y_2)=E(f)E(g)=a^2.$$

Пример 2. Если $0<a<1$, то случайные переменные $f=s_1$ и $g=s=s_1+s_2$ *зависимы*. Для них

$$E(f\cdot g)=E\left(s_1^2+s_1s_2\right)=E\left(s_1^2\right)+E(s_1s_2)=a+a^2\neq a\cdot 2a=E(f)\,E(g).$$

Пример 3. Если $0<a<1$, то случайные переменные $f=s_1-s_2$ и $g=s_1+s_2$ *зависимы:*

$$P(f^{-1}(0)g^{-1}(0))=P(\{00\})=(1-a)^2\neq[a^2+(1-a)^2](1-a)^2=$$
$$=P(\{11,\ 00\})P(\{00\})=P(f^{-1}(0))P(g^{-1}(0)).$$

В то же время для них

$$E(f\cdot g)=E[(s_1-s_2)(s_1+s_2)]=E\left(s_1^2-s_2^2\right)=$$
$$=E\left(s_1^2\right)-E\left(s_2^2\right)=a-a=0=E(s_1)-E(s_2)=$$
$$=E(s_1-s_2)=E(s_1-s_2)\,E(s_1+s_2)=E(f)\,E(g).$$

Замечание. Примеры 2 и 3 показывают, что правило умножения для среднего для некоторых зависимых случайных переменных верно, а для некоторых — нет.

Пример 4. Случайные переменные $f_1=i_{c_1}$, $f_2=i_{c_2}$, $f_3=i_{c_3}$ для примера Бернштейна с раскрашенным тетраэдром, рассмотренного в пункте 5 из § 5 главы 3 части I, зависимы. Для них

$$E(f_1f_2f_3)=P(c_1c_2c_3)=1/4\neq 1/8=P(c_1)P(c_2)P(c_3)=E(f_1)E(f_2)E(f_3).$$

Пример 5. Для модели Бернулли n испытаний с вероятностью успеха a рассмотрим *непересекающиеся* части K и L множества $\{1,\ldots,n\}$, составленные соответственно из k и l номеров. Положим $f_j=i_{Y_j}$ для каждого номера $j\in K$ и $f_j=i_{H_j}$ для каждого номера $j\in L$. Рассмотрим семейство $(f_j)_{j\in I}$ случайных переменных f_j с множеством индексов $I=K+L$. Из независимости испытаний в модели Бернулли следует независимость случайных переменных f_j этого семейства. Из правила индикатора, общего правила умножения и равенств

$$P(Y_j)=a,\quad P(H_j)=1-a$$

следует формула успехов и неудач:

$$P\left(\bigcap_{j\in K} У_j \bigcap_{j\in L} H_j\right) = E\left(\prod_{j\in I} f_j\right) = \prod_{j\in I} E(f_j) =$$
$$= \prod_{j\in K} P(У_j)\, P(H_j) = a^k (1-a)^l.$$

Таким образом, если доказать равенства для вероятностей успеха и неудачи, не используя формулу успехов и неудач, то эта формула будет следствием общего правила умножения для среднего.

В частности, среднее произведения

$$P = \prod_{1\leqslant j\leqslant n} s_j$$

выражается равенством

$$E(p) = a^n.$$

Задачи к § 2 собраны в § 1 главы 2 части III.

§ 3. ДИСПЕРСИЯ

Среднее является очень грубой характеристикой случайной переменной: ее значения могут сильно отклоняться от среднего. Поэтому — для более точного описания случайной переменной — рассматривается ее *дисперсия*, измеряющая близость значений случайной переменной к среднему.

1.3. Примеры

Рассмотрим несколько примеров, разъясняющих понятие дисперсии.

Пример 1. В примере 1 из § 1 игры с костью среднее выигрыша f равно 0 при любой ставке c. Чем больше c, тем сильнее значения $-c$ и $+c$ случайной переменной f отклоняются от среднего $E(f)=0$ (рис. 31).

$-c$ -2 -1 0 1 2 c

Рис. 31.

Это отклонение можно описать средним $D(f)$ случайной переменной $(f-\mathrm{E}(f))^2$, которое называется *средним квадратичным отклонением* или *дисперсией* случайной переменной f:

$$D(f) = E(c^2) = c^2.$$

Механически дисперсию $D(f)$ в рассматриваемом примере можно представить себе как *момент инерции относительно центра* гантели длины $2c$ с массами $1/2$ на концах: чем длиннее гантель, тем больше момент ее инерции.

Удобно использовать также стандартное отклонение

$$S(f)=\sqrt{D(f)}=c.$$

В данном случае абсолютные отклонения всех значений случайной переменной f от среднего $E(f)=0$ *не превосходят стандартное*:

$$|f(u)-E(f)|=c=S(f).$$

Пример 2. В примере 2 из § 1 игры с монетой среднее $E(f)$ выигрыша f равно 0. Значения -10 и $+10$ случайной переменной f отклоняются от среднего $E(f)=0$ (рис. 32).

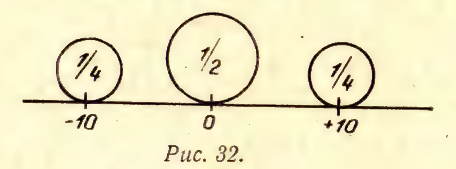


Рис. 32.

Это отклонение можно описать средним $D(f)$ случайной переменной $(f-E(f))^2$, называемым дисперсией случайной переменной f:

$$D(f)=E[(f-E(f))^2]=(-10-0)^2 1/4+(0-0)^2 1/2+$$
$$+(+10-0)^2 1/4=50.$$

Удобно также рассматривать *стандартное отклонение* случайной переменной f:

$$S(f)=\sqrt{D(f)}=\sqrt{50}\approx 7.$$

В этом примере абсолютные отклонения значений -10 и $+10$ случайной переменной f от среднего $E(f)$ *превосходят стандартное отклонение* $s(f)=\sqrt{50}$, но все значения f не превосходят $(1{,}5)s(f)$. В пределах стандартного отклонения расположена *половина* распределения случайной переменной f:

$$P(\{u:|f(u)-E(f)|\leqslant S(f)\})=P\{0\}=1/2.$$

Пример 3. В примере 3 из § 1 с выстрелом по мишени среднее $E(f)$ числа очков f равно 8,6. Значения случайной переменной f отклоняются от среднего $E(f)=8{,}6$. Это отклонение можно описать средним $D(f)$ случайной переменной $(f-E(f))^2$, т. е. *дисперсией* случайной переменной f:

$$D(f)=E[(f-E(f))^2]=(0-8{,}6)^2 0{,}01+(1-8{,}6)^2 0{,}01+(2-8{,}6)^2 0{,}01+$$
$$+(3-8{,}6)^2 0{,}01+(4-8{,}6)^2 0{,}01+(5-8{,}6)^2 0{,}05+(6-8{,}6)^2 0{,}05+$$
$$+(7-8{,}6)^2 0{,}05+(8-8{,}6)^2 0{,}10+(9-8{,}6)^2 0{,}20+(10-8{,}6)^2 0{,}50\approx 4{,}4$$

Удобно использовать также *стандартное отклонение*

$$S(f)=\sqrt{D(f)}=\sqrt{4{,}4}\approx 2{,}1.$$

В этом примере абсолютные отклонения значений случайной переменной f от среднего $E(f)=8{,}6$ не превышают 4,5 стандартных отклонений:

$$|f(u)-E(f)| \leqslant (4{,}5)S(f)$$

для каждого исхода u. В пределах стандартного отклонения расположено 85% распределения случайной переменной f:

$$P(\{:u \mid |f(u)-E(f)| \leqslant S(f)\})=P(\{7,\ 8,\ 9,\ 10\})=0{,}85.$$

В пределах *двух* стандартных отклонений расположено уже 95% распределения f:

$$P(\{u: |f(u)-E(f)| \leqslant 2S(f)\})=P(\{5,\ 6,\ 7,\ 8,\ 9,\ 10\})=0{,}95.$$

2.3. Определения

Рассмотрим конечную вероятностную модель $(U, \mathscr{A};\ p,\ P)$, случайную переменную f на U, ее множество значений X, элементарное распределение q и среднее $E(f)$.

Определение 2. *Дисперсией случайной переменной f называется число*

$$D(f)=E[(f-E(f))^2].$$

Обозначим множество всех случайных переменных на U буквой $\mathscr{F}$. Функция D на множестве $\mathscr{F}$, значением которой для каждой случайной переменной f из $\mathscr{F}$ является ее дисперсия $D(f)$, называется *дисперсией.*

При вычислении дисперсии часто бывает полезно

Предложение 2. $D(f)=\Sigma(x-E(f))^2 q(x)$.

Доказательство. Используя определение среднего и группируя слагаемые по аналогии с доказательством предложения 1, получаем:

$$D(f)=\sum_{u\in U}(f(u)-E(f))^2 p(u)=\sum_{x\in X}\Bigl(\sum_{u\in f^{-1}(x)}(f(x)-E(f))^2 p(u)\Bigr)=$$

$$=\sum_{x\in X}\Bigl(\sum_{u\in f^{-1}(x)}(x-E(f))^2 p(u)\Bigr)=\sum_{x\in X}\Bigl((x-E(f))^2\sum_{u\in f^{-1}(x)}p(u)=$$

$$=\sum_{x\in U}(x-E(f)\Bigr)^2 P(f^{-1}(x))=\sum_{x\in X}(x-E(f))^2 q(x).$$

Предложение 2 доказано.

Пример. В примере о блуждании мухи по решетке (пункт 1 из § 2), используя предложение 2, получаем:

$$D(d)=[2\sqrt{2}-(1+\sqrt{2}/2)]^2 1/4+[2-(1+\sqrt{2}/2)]^2 1/2+$$

$$+[0-(1+\sqrt{2}/2)]^2 1/4=5/2-\sqrt{2}.$$

Пример. Для индикатора $f=i_A$ произвольного события A из $\mathscr{A}$ имеем:

$$X=\{1, 0\}; \quad q(1)=P(A), \quad q(0)=1-P(A); \quad E(f)=P(A);$$
$$D(f)=(1-P(A))^2P(A)+(0-P(A))^2(1-P(A))=$$
$$=P(A)(1-P(A)+P(A))=P(A)(1-P(A)).$$

Таким образом, для дисперсии D верно

Правило индикатора:

$$D(i_A)=P(A)(1-P(A)).$$

В частности, если $P(A)=1/2$, то $D(i_A)=1/4$.

Упражнение. Доказать, что

$$D(i_A)\leqslant 1/4$$

для каждого события A.

Так как случайные переменные с одинаковыми распределениями имеют одинаковые средние, то из предложения 2 вытекает

Следствие. *Если случайные переменные имеют одинаковые распределения, то их дисперсии равны.*

Это следствие показывает, что дисперсия, как и среднее, определяется *распределением* случайной переменной.

Определение 2′. *Стандартным отклонением случайной переменной f называется число*

$$S(f)=\sqrt{D(f)}.$$

Пример. $S(i_A)=\sqrt{P(A)(1-P(A))}.$

В частности, если $P(A)=1/2$, то $S(i_A)=1/2$.

Функция S на множестве $\mathscr{F}$, значением которой для каждой случайной переменной f из $\mathscr{F}$ является ее стандартное отклонение $S(f)$, называется *стандартным отклонением.*

3.3. Содержание и механическая интерпретация

Дисперсию $D(f)$ случайной переменной f с содержательной точки зрения можно считать *оценкой отклонения* значения случайной переменной f от ее ожидаемого значения $E(f)$. Примеры 1—3 пункта 1 поясняют это.

В механической системе точек x с массами $q(x)$ дисперсия $D(f)$ является *моментом инерции относительно центра масс* $E(f)$.

4.3. Свойства дисперсии

Положительность. $D(f)\geqslant 0$.

Доказательство. Из положительности среднего E и квадрата $(f-E(f))^2$ следует, что

$$D(f)=E[(f-E(f))^2]\geqslant 0.$$

Рассмотрим произвольную постоянную c на U.

Правило постоянной:

$$D(c)=0.$$

Доказательство. Из правила постоянной для среднего E следует, что

$$D(c)=E[(c-E(c))^2]=E[(c-c)^2]=E(0)=0.$$

Общее правило постоянной. *Дисперсия случайной переменной равна нулю, если и только если f почти всюду равна некоторой постоянной c.*

Доказательство. Если $f=c$ почти всюду, то $E(f)=c$, $f-E(f)=0$ почти всюду и

$$D(f)=E[(f-E(f))^2]=0.$$

Обратно. Если

$$D(f)=\Sigma(f(u)-E(f))^2p(u)=0,$$

то $p(u)=0$ для каждого исхода u, для которого $f(u)\neq E(f)$, и, значит, $f=E(f)$ почти всюду.

Рассмотрим произвольное число a.

Правило умножения на число:

$$D(af)=a^2D(f).$$

Доказательство. Из правила умножения на число для среднего E следует, что

$$D(af)=E[(af-E(af))^2]=E[(af-aE(f))^2]=E[a^2(f-E(f))^2]=$$
$$=a^2E[(f-E(f))^2]=a^2D(f).$$

Правило прибавления постоянной:

$$D(f+c)=D(f).$$

Доказательство. Из правила сложения и правила постоянной для среднего E следует, что

$$D(f+c)=E[((f+c)-E(f+c))^2]=E[(f+c-E(f)-E(c))^2]=$$
$$=E[(f+c-E(f)-c)^2]=E[(f-E(f))^2]=D(f).$$

Рассмотрим *независимые* случайные переменные f и g на U.

Правило сложения:

$$D(f+g)=D(f)+D(g).$$

Доказательство. Из линейности среднего E следует, что

$$D(f+g)=E[((f+g)-E(f+g))^2]=E[(f+g-E(f)-E(g))^2]=$$
$$=E[((f-E(f))+(g-E(g)))^2]=E[(f-E(f))^2+(g-E(g))^2+$$
$$+2(f-E(f))(g-E(g))]=E[(f-E(f))^2]+E[(g-E(g))^2]+$$
$$2E[(f-E(f))(g-E(g))]=D(f)+D(g)+2E[(f-E(f))(g-E(g))].$$

Снова используя линейность и правило постоянной для среднего E, правило умножения для независимых случайных переменных f и g, получаем:

$$E[(f-E(f))(g-E(g))]=E[f\cdot g-fE(g)-E(f)g+E(f)E(g)]=$$
$$=E(f\cdot g)-E[fE(g)]-E[E(f)g]+E[E(f)E(g)]=E(f)E(g)-$$
$$-E(f)E(g)-E(f)E(g)+E(f)E(g)=0.$$

Таким образом,

$$D(f+g)=D(f)+D(g).$$

Правило сложения для дисперсий доказано.

Рассмотрим семейство $(f_j)_{j\in I}$ *попарно независимых* случайных переменных f_j на U.

Общее правило сложения:

$$D(\Sigma f_j)=\Sigma D(f_j).$$

Доказательство. Используя общее правило сложения для среднего E, получаем:

$$D\left(\sum_{j\in I} f_j\right)=E\left[\left(\sum_{j\in I} f_j-E\left(\sum_{j\in I} f_j\right)\right)^2\right]=E\left[\left(\sum_{j\in I} f_j-\sum_{j\in I} E(f_j)\right)^2\right]=$$
$$=E\left[\left(\sum_{j\in I}(f_j-E(f_j))\right)^2\right]=E\left[\sum_{i\in I}(f_i-E(f_i))\sum_{j\in I}(f_j-E(f_j))\right]=$$
$$=E\left[\sum_{i\in I}\sum_{j\in I}(f_i-E(f_i))(f_j-E(f_j))\right]=$$
$$=\sum_{i\in I}\sum_{j\in I}E[(f_i-E(f_i))(f_j-E(f_j))].$$

Если $i=j$, то

$$E[(f_j-E(f_j))(f_j-E(f_j))]=E[(f_j-E(f_j))^2]=D(f_j).$$

Если $i\neq j$, то вследствие независимости случайных переменных $f=f_i$ и $g=f_j$

$$E[(f_i-E(f_i))(f_j-E(f_j))]=0.$$

Таким образом,

$$D(\Sigma f_j)=\Sigma D(f_j).$$

Общее правило сложения для дисперсии доказано.

При вычислении дисперсии часто бывает полезна

Формула квадратов:

$$D(f)=E(f^2)-E^2(f).$$

Доказательство. Из линейности среднего E и правила постоянной следует, что

$$D(f)=E[(f-E(f))^2]=E[f^2-2fE(f)+E^2(f)]=E(f^2)-2E(f)E(f)+$$
$$E[E^2(f)]=E(f^2)-2E^2(f)+E^2(f)=E(f^2)-E^2(f).$$

Пример. В примере о блуждании мухи по решетке (пункт 1 из § 2), используя формулу квадратов, получаем:

$$D(d)=E(d^2)-E^2(d)=[2^2\cdot 1/4+(2\sqrt{2})^2 1/4+2^2 1/4]-[1+\sqrt{2}/2]^2=$$
$$=5/2-\sqrt{2}.$$

Результат совпадает с полученным в пункте 2.

5.3. Примеры

Рассмотрим несколько примеров, разъясняющих использование свойств дисперсии.

В примерах 1—3 используется *модель Бернулли $n=2$ испытаний с вероятностью успеха a.*

Пример 1. Случайные переменные $f=s_1=i_{У_1}$ и $g=s_2=i_{У_2}$ *независимы.* Для них

$$D(f+g)=D(f)+D(g)=2a(1-a).$$

Пример 2. Если $0<a<1$, то случайные переменные $f=s_1$ и $g=s_1+s_2$ *зависимы.* Для них

$$D(f+g)=D(2s_1+s_2)=D(2s_1)+D(s_2)=4D(s_1)+D(s_2)=5a(1-a)=$$
$$\neq 3a(1-a)=D(s_1)+D(s_1+s_2)=D(f)+D(g).$$

Пример 3. Если $0<a<1$, то случайные переменные $f=s_1-s_2$ и $g=s_1+s_2$ *зависимы.* В то же время для них

$$D(f+g)=D(2s_1)=4D(s_1)=4a(1-a)=D(s_1)+D(-s_2)+D(s_1)+$$
$$+D(s_2)=D(s_1-s_2)+D(s_1+s_2)=D(f)+D(g).$$

Замечание. Примеры 2 и 3 показывают, что правило сложения для дисперсии для некоторых зависимых случайных переменных верно, а для некоторых — нет.

Пример 4. Случайные переменные $f_1=i_{c_1}$, $f_2=i_{c_2}$, $f_3=i_{c_3}$ для примера Бернштейна с раскрашенным тетраэдром попарно независимы. Для них

$$D(f_1+f_2+f_3)=D(f_1)+D(f_2)+D(f_3).$$

Пример 5. Рассмотрим *модель Бернулли n испытаний с вероятностью успеха a* и случайную переменную

$$s=\sum_{1\leqslant j\leqslant n} s_j=\sum_{1\leqslant j\leqslant n} i_{У_j},$$

описывающую *число успехов.*

Случайные переменные s_j являются независимыми индикаторами. Используя общее правило сложения и правило индикатора для дисперсии, получаем:

$$D(s)=\sum_{1\leqslant j\leqslant n} D(s_j)=\sum_{1\leqslant j\leqslant n} P(У_j)(1-P(У_j))=n\cdot a(1-a).$$

В частности, если $n=100$ и $a=1/2$, то $D(s)=25$ и $S(s)=$ $=\sqrt{D(s)}=5$.

Пример 6. Для *модели Бернулли n испытаний с вероятностью успеха a* рассмотрим случайную переменную s/n, описывающую *частоту успеха*. Используя результат примера 5 и правило умножения на число для дисперсии, получаем:

$$D(s/n)=D(s)/n^2=na(1-a)/n^2=a(1-a)/n.$$

В частности, если $n=100$ и $a=1/2$, то

$$D(s/n)=1/400 \text{ и } S(s/n)=\sqrt{D(s/n)}=1/20.$$

Задачи к § 3 собраны в § 1 главы 2 части III.

§ 4*. КОРРЕЛЯЦИЯ

Коэффициент корреляции случайных переменных позволяет определить, являются они *линейно зависимыми* или нет: если абсолютное значение коэффициента корреляции равно единице, то случайные переменные линейно зависимы, если оно строго меньше единицы,— то нет. Можно сказать, что коэффициент корреляции измеряет близость зависимости случайных переменных к линейной.

Общая задача о зависимости или независимости каждых двух случайных переменных с помощью коэффициента корреляции решается лишь частично: если коэффициент корреляции не равен нулю, то случайные переменные зависимы, но если он равен нулю, то они могут быть как зависимы, так и независимы.

1.4. Примеры

Рассмотрим несколько примеров, поясняющих понятие *корреляционной зависимости* случайных переменных.

Коэффициент корреляции $K(A, B)$ возможных событий A и B определяется равенством

$$K(A, B)=\frac{P(AB)-P(A)P(B)}{\sqrt{P(A)(1-P(A))P(B)(1-P(B))}}.$$

Естественно назвать *коэффициентом корреляции* $K(f, g)$ *индикаторов* $f=i_A$ и $g=i_B$ событий A и B коэффициент корреляции $K(A, B)$ этих событий. По правилам индикаторов, для среднего E и дисперсии D для индикаторов $f=i_A$ и $g=i_B$ событий A и B верны равенства

$$E(fg)=P(AB), \quad E(f)E(g)=P(A)P(B),$$
$$D(f)=P(A)(1-P(A)), \ D(g)=P(B)(1-P(B)).$$

Следовательно,

$$K(f, g) = \frac{E(f \cdot g) - E(f)E(g)}{\sqrt{D(f)D(g)}}.$$

Этим равенством можно определить коэффициент корреляции каждых *невырожденных* случайных переменных f и g, для которых $D(f) \neq 0$ и $D(g) \neq 0$.

В примерах 1—5 используется *модель Бернулли $n=2$ испытаний с вероятностью успеха a* $(0<a<1)$.

Пример 1. Случайные переменные $f = s_1 = i_{y_j}$ и $g = s_2 = i_{y_k}$ *независимы*. Для них

$$E(fg) = E(f)E(g), \quad K(f, g) = 0.$$

Пример 2. Случайные переменные $f = s_1$ и $g = s_1 + s_2$ зависимы. Для них

$$K(f, g) = \frac{a + a^2 - 2a^2}{\sqrt{a(1-a)2a(1-a)}} = \frac{1}{\sqrt{2}}.$$

Замечание. Примеры 2 и 3 показывают, что коэффициент корреляции для некоторых зависимых случайных переменных равен нулю, а для некоторых — нет. Это отличает его от коэффициента корреляции событий.

Пример 4. Случайные переменные $f = s_1$ и $g = 2s_1 + 3$ линейно зависимы. Для них

$$K(f, g) = \frac{5a - a(2a+3)}{\sqrt{a(1-a)4a(1-a)}} = +1.$$

Пример 5. Случайные переменные $f = s_1$ и $g = -s_1 + 5$ линейно зависимы. Для них

$$K(f, g) = \frac{3a - a(-2a+5)}{\sqrt{a(1-a)4a(1-a)}} = -1.$$

2.4. Определение

Рассмотрим конечную вероятностную модель $(U, \mathcal{A}; p, P)$ и случайные переменные f, g с *невырожденными* распределениями:

$$D(f) > 0, \ D(g) > 0.$$

Определение 3. *Коэффициентом корреляции случайных переменных f и g называется число*

$$K(f, g) = \frac{E(f \cdot g) - E(f)E(g)}{\sqrt{D(f)D(g)}}.$$

Коэффициент корреляции *симметричен*:

$$K(f, g) = K(g, f).$$

Среди свойств коэффициента корреляции выделяется

Основное свойство. *Если случайные переменные f и g зависимы, то $K(f, g) = 0$.*

Доказательство. Если случайные переменные f и g независимы, то по правилу умножения для среднего

$$E(f\cdot g)=E(f)E(g)$$

и, следовательно,

$$E(f\cdot g)-E(f)E(g)=0,\ K(f,\ g)=0.$$

Замечание. Пример 3 пункта 1 показывает, что обратное утверждение не верно: из равенства нулю коэффициента корреляции *не следует* независимость случайных переменных. Условие $K(f,\ g)=0$ является необходимым, но не достаточным условием независимости случайных переменных f и g.

Основное свойство коэффициента корреляции можно сформулировать также в следующей эквивалентной форме: *если $K(f,\ g)\neq 0$, то случайные переменные f и g зависимы.* Условие $K(f,\ g)\neq 0$ является достаточным, но не необходимым условием зависимости случайных переменных f и g.

При исследовании свойств коэффициента корреляции бывает полезно

Предложение 3. $K(f,\ g)=\dfrac{E[(f-E(f))(g-E(g))]}{S(f-E(f))S(g-E(g))}$.

Доказательство. В самом деле,

$$E[(f-E(f))\ (g-E(g))]=E[(f\cdot g-f\cdot E(g)-E(f)\cdot g+E(f)E(g)]=$$
$$=E(f\cdot g)-E(f)E(g)-E(f)E(g)+E(f)E(g)=E(f\cdot g)-E(f)E(g),$$
$$S(f-E(f)=\sqrt{D(f-E(f))}=\sqrt{D(f)},\ S(g-E(g))=\sqrt{D(g-E(g))}=$$
$$=\sqrt{D(g)}.$$

Рассмотрим произвольные постоянные b и d. Из предложения 3 вытекает

Следствие. $K(f+b,\ g+d)=K(f,\ g)$.

Доказательство. В самом деле,

$$(f+b)-E(f+b)=f-E(f),$$
$$g+d-E(g+d)=g-E(g).$$

3.4. Геометрическая интерпретация

По-прежнему рассматриваются случайные переменные на данном множестве исходов U.

Каждую случайную переменную f можно центрировать с помощью ее среднего $E(f)$ и рассматривать случайную переменную

$$f_0=f-E(f),$$

среднее которой равно нулю:

$$E(f_0)=E(f)-E[E(f)]=E(f)-E(f)=0.$$

Рассмотрим *модель Бернулли $n=2$ испытаний с вероятностью успеха a* $(0<a<1)$.

Пример 1. Если $f=s_1$ и $g=s_2$, то $f_0=s_1-a$ и $g_0=s_2-a$.

Пример 2. Если $g=s_1+s_2$, то $g_0=s_1+s_2-2a=(s_1-a)+(s_2-a)$.

Пример 3. Если $f=s_1-s_2$, то $f_0=f$.

Условимся каждую случайную переменную, среднее которой равно нулю, называть *центрированной*. Из линейности среднего следует, что сумма и произведение на число каждых центрированных случайных переменных являются центрированными случайными переменными:

$$E(f_0+g_0)=E(f_0)+E(g_0)=0+0=0,$$
$$E(cf_0)=cE(f_0)=c\cdot 0=0,$$

если

$$E(f_0)=E(g_0)=0.$$

Поэтому центрированные случайные переменные можно представлять себе как *векторы*.

Стандартное отклонение $S(f_0)$ можно интерпретировать как *длину* центрированной случайной переменной f_0.

Из определения стандартного отклонения следует, что

$$S(f_0)\geqslant 0.$$

Определение. *Случайная переменная f называется вырожденной, если и только если она имеет вырожденное распределение.*

Из общего правила постоянной для дисперсии следует, что случайная переменная f является вырожденной, если и только если она *почти всюду равна некоторой постоянной c.*

Равенство $S(f_0)=0$ эквивалентно вырожденности f_0.

При исследовании дальнейших свойств стандартного отклонения будет использоваться неравенство Шварца для среднего и стандартного отклонения.

Рассмотрим произвольные случайные переменные f и g.

Определение. *Случайные переменные f и g называются линейно зависимыми, если и только если существуют число a и постоянная b такие, что $a\cdot f-g=b$ почти всюду.*

Примеры. В примерах 4, 5 пункта 1 случайные переменные f и g *линейно зависимы.* В примерах 1—3 — *нет.*

Рассмотрим произвольные *центрированные* случайные переменные f_0 и g_0.

Неравенство Шварца: $|E(f_0\cdot g_0)|\leqslant S(f_0)S(g_0)$, *причем равенство верно, если и только если f_0 и g_0 линейно зависимы.*

Доказательство. Так как случайные переменные f_0 и g_0 центрированные, то

$$S(f_0)=\sqrt{E(f_0^2)},\quad S(g_0)=\sqrt{E(g_0^2)}$$

и, следовательно, доказываемое неравенство эквивалентно неравенству

$$d=E^2(f_0\cdot g_0)-E(f_0^2)\,E(g_0^2)\leqslant 0.$$

Левая часть d этого неравенства является дискриминантом квадратного уравнения

$$E\left(f_0^2\right)x^2-2E\left(f_0\cdot g_0\right)x+E\left(g_0^2\right)=0.$$

Из линейности и положительности среднего E вытекает, что

$$E\left(f_0^2\right)x^2-2E\left(f_0\cdot g_0\right)x+E\left(g_0^2\right)=E[(f_0x-g_0)^2]\geqslant 0.$$

Поэтому рассматриваемое уравнение либо имеет корень

$$x=a=E\left(f_0\cdot g_0\right)/E\left(f_0^2\right),$$

либо не имеет вещественных корней.

Соответственно либо $d=0$, либо $d<0$. Следовательно, всегда $d\leqslant 0$, причем равенство $d=0$ верно, если и только если

$$E[(af_0-g_0)^2]=0.$$

В свою очередь, это равенство верно, если и только если $af_0-g_0=0$ почти всюду. Лемма 1 доказана.

Из неравенства Шварца вытекает *неравенство треугольника* и *правило умножения на число* для стандартного отклонения. Аналогичные неравенство и правило верны для *длины вектора*.

Рассмотрим произвольные *центрированные* переменные f_0, g_0 и число c.

Неравенство треугольника: $S(f_0+g_0)\leqslant S(f_0)+S(g_0)$.

Доказательство. Используя линейность среднего E и неравенство Шварца, получаем:

$$S^2(f_0+g_0)=E[(f_0+g_0)^2]=E\left[f_0^2+2f_0\cdot g_0+g_0^2\right]=$$
$$=E\left(f_0^2\right)+E\left(g_0^2\right)+2E\left(f_0\cdot g_0\right)=S^2(f_0)+S^2(g_0)+2E\left(f_0\cdot g_0\right)\leqslant$$
$$\leqslant S^2(f_0)+S^2(g_0)+2S(f_0)\,S(g_0)=[S(f_0)+S(g_0)]^2.$$

Отсюда следует неравенство треугольника.

Правило умножения на число:

$$S(cf_0)=|c|\cdot S(f_0).$$

Доказательство. Используя правило умножения на число для среднего E, получаем:

$$S(cf_0)=\sqrt{E[(cf_0)^2]}=\sqrt{c^2E\left(f_0^2\right)}=|c|\sqrt{E\left(f_0^2\right)}=|c|\cdot S(f_0).$$

Коэффициент корреляции $K(f_0, g_0)$ можно интерпретировать как *косинус угла* между невырожденными центрированными случайными переменными f_0 и g_0:

$$S(f_0)>0,\ S(g_0)>0.$$

Из неравенства Шварца вытекает

Правило косинуса:

$$|K(f_0, g_0)|\leqslant 1,$$

причем равенство верно, если и только если f_0 и g_0 линейно зависимы.

Доказательство. Так как случайные переменные f_0 и g_0 центрированные, то

$$K(f_0, g_0) = \frac{E(f_0 \cdot g_0)}{S(f_0)\, S(g_0)}.$$

Отсюда и из неравенства Шварца следуют доказываемые неравенство и утверждение.

Правило косинуса позволяет интерпретировать равенство $K(f_0, g_0)$ как *ортогональность* случайных переменных f_0 и g_0, а равенство $|K(f_0, g_0)| = 1$ — как их *параллельность*. Если $K(f_0, g_0) = +1$, то f_0 и g_0 *одинаково направлены*, а если $K(f_0, g_0) = -1$ — *противоположно*. Можно также сказать, что $|K(f_0, g_0)| = 1$, если и только если f_0 и g_0 *лежат на одной прямой*.

Все эти термины употребляются для *векторов на плоскости* и *косинуса угла* между ними.

Сказанное можно пояснить рисунками для примеров 1—5 пункта 1 (рис. 33).

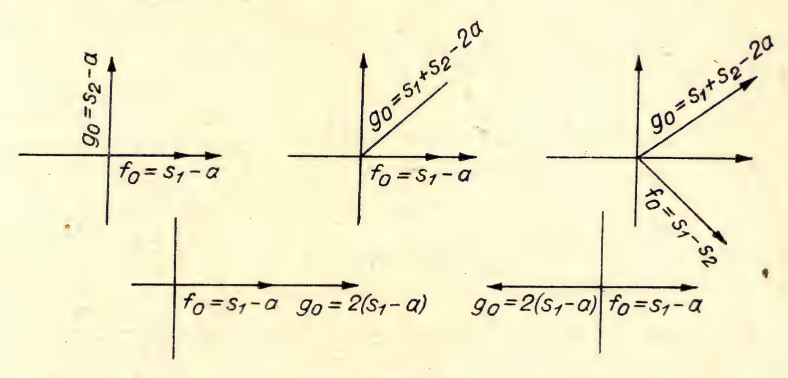


Рис. 33.

4.4. Свойства

Рассмотрим *невырожденные* случайные переменные f, g и определяемые ими *центрированные* случайные переменные

$$f_0 = f - E(f),\quad g_0 = g - E(g).$$

Из предложения 3 следует, что

$$K(f, g) = K(f_0, g_0).$$

Поэтому все сказанное в пункте 3 для коэффициента корреляции центрированных случайных переменных $K(f_0, g_0)$ верно и для $K(f, g)$.

Правило косинуса: $|K(f, g)| \leqslant 1$, *причем равенство верно, если и только если f и g линейно зависимы.*

Доказательство. По уже доказанному правилу косинуса для центрированных случайных переменных

$$|K(f, g)| = |K(f_0, g_0)| \leqslant 1.$$

Причем равенство верно, если и только если существует число a такое, что $af_0 - g_0 = 0$ почти всюду.
Для каждого исхода u равенство

$$af_0(u) - g_0(u) = 0$$

эквивалентно равенству

$$af(u) - g(u) = b,$$

где

$$b = aE(f) - E(g).$$

Следовательно, равенство $af_0 - g_0 = 0$ (почти всюду) эквивалентно равенству $af - g = b$ (почти всюду).

Вместе с невырожденными случайными переменными f, g рассмотрим также числа $a \neq 0$, $c \neq 0$ и постоянные b, d. Говорят, что случайные переменные $af + b$ и $cd + d$ получаются из случайных переменных f и g *линейными преобразованиями.*

Для каждого вещественного числа x определим знак $\operatorname{sgn} x$ числа x равенством

$$\operatorname{sgn} x = \begin{cases} +1 & (x > 0), \\ 0 & (x = 0), \\ -1 & (x < 0). \end{cases}$$

Правило линейных преобразований:

$$K(af + b, cg + d) = \operatorname{sgn}(ab) K(f, g).$$

Доказательство. В самом деле, используя линейность среднего E, правила умножения на число и прибавления постоянной для дисперсии D, получаем:

$$E[(af + b)(cg + d)] - E(af + b) E(cg + d) = ac[E(f \cdot g) - E(f) E(g)],$$

$$S(af + b) = |a| \cdot S(f), \quad S(cg + d) = |c| \cdot S(g).$$

Следовательно,

$$K(af + b, \ cg + d) = \frac{ac}{|ac|} K(f, g) = \operatorname{sgn}(ab) K(f, g).$$

Следствие. $K(-f, g) = K(f, -g) = -K(f, g)$.

Примеры 4, 5 пункта 1 поясняют применение правила линейных преобразований для коэффициента корреляции.

5.4. Примеры

Рассмотрим несколько примеров вычисления коэффициента корреляции.

Пример 1. Из формулы квадратов для дисперсии следует, что

$$K(f, f) = \frac{E(f^2) - E^2(f)}{D(f)} = 1$$

для каждой *невырожденной* случайной переменной f. Тот же результат можно получить, используя предложение 3 и правило прибавления постоянной для дисперсии:

$$K(f, f) = \frac{E[(f - E(f))^2]}{S^2(f - E(f))} = \frac{D(f)}{D(f)} = 1$$

Пример 2. Из равенства $K(f, f) = 1$ и правила линейных преобразований следует, что

$$K(f, -f) = -1.$$

В примерах 3—5 используется *модель Бернулли n испытаний с вероятностью успеха a* $(0 < a < 1)$.

Пример 3. Из результата примера 1 и правила линейных преобразований следует, что коэффициент корреляции числа успехов s и частоты успехов s/n выражается равенствами:

$$K(s, s/n) = K(s, s) = 1.$$

Пример 4. Так как

$$E(s_1 s) = E\left(\sum_{1 \leqslant j \leqslant n} s_1 s_j\right) = \sum_{1 \leqslant j \leqslant n} E(s_1 s_j) = a + (n - 1) a^2,$$

$$E(s_1) = a, \quad E(s) = na, \quad D(s_1) = a(1 - a), \quad D(s) = na(1 - a),$$

то

$$K(s_1, s) = \frac{a + (n-1) a^2 - na^2}{\sqrt{n}\, a(1 - a)} = \frac{1}{\sqrt{n}}.$$

Замечание. Результат соответствует интуитивному представлению о том, что с увеличением числа n слагаемых сумма s должна меньше зависеть от одного из них.

Пример 5. Рассмотрим случайные переменные

$$p = \prod_{1 \leqslant j \leqslant n} s_j, \quad s = \sum_{1 \leqslant j \leqslant n} s_j.$$

Как было показано,

$$E(s) = na, \; D(s) = na(1 - a).$$

Так как случайные переменные s_j семейства $(s_j)_{1 \leqslant j \leqslant n}$ *независимы*, то из общего правила умножения следует, что

$$E(p) = \prod_{1 \leqslant j \leqslant n} E s_j = a^n.$$

Используя равенство $p^2 = p$ и формулу квадратов для дисперсии, получаем:

$$D(p) = E(p^2) - E^2(p) = E(p) - E^2(p) = a^n (1 - a^n).$$

Наконец, так как $p \cdot s = n \cdot p$, то

$$E(p \cdot s) = E(n \cdot p) = nE(p) = n \cdot a^n.$$

Следовательно,

$$K(p, s) = \frac{na^n - na^n a}{\sqrt{na(1-a)a^n(1-a^n)}} = \sqrt{na^{n-1}\frac{1-a}{1-a^n}}.$$

В частности, если $n=1$, то $K(p, s) = 1$, а если $n=2$, то

$$K(p, s) = \sqrt{\frac{2a}{1+a}}.$$

Упражнение. Доказать, что $K(p, s)$ уменьшается с увеличением n, и объяснить это.

§ 5*. ИНФОРМАЦИЯ

Термин *информация* используется здесь в очень узком и точно определенном смысле. Информация случайных переменных является удобной *мерой их зависимости*. С ее помощью полностью решается общая задача о зависимости или независимости каждых двух случайных переменных: если информация не равна нулю, то случайные переменные зависимы, если равна нулю, то они независимы.

Информация определяется как среднее некоторой специально выбранной случайной переменной.

1.5. Примеры

Рассмотрим несколько примеров, поясняющих используемое понятие информации.

Рассмотрим *модель Бернулли $n=2$ испытаний с вероятностью успеха $a=1/2$*. Как в пункте 4 из § 4 главы 1, для каждых случайных переменных f и g на множестве исходов $U=\{11, 10, 01, 00\}$ определим случайную переменную:

$$i(f, g) = \log\frac{s(f(11), g(11))}{q(f(11))\, r(g(11))} i_{11} + \log\frac{s(f(10), g(10))}{q(f(10))\, r(g(10))} i_{10} +$$

$$+ \log\frac{s(f(01), g(01)}{q(f(01))\, r(g(01)} + \log\frac{s(f(00), g(00))}{q(f(00))\, r(g(00))} i_{00}$$

и рассмотрим ее среднее:

$$I(f, g) = E[i(f, g)] =$$

$$= P(11)\log\frac{s(f(11), g(11))}{q(f(11)\, r(g(11))} + p(10)\log\frac{s(f(10), g(10))}{q(f(10))\, r(g(10))} +$$

$$+ p(01)\log\frac{s(f(01), g(01)}{q(f(01))\, r(g(01))} + p(00)\log\frac{s(f(00)), g(00))}{q(f(00))\, r(g(00))}.$$

Замечание. Здесь, как и всюду в книге, рассматриваются логарифмы с основаниями *строго большими единицы.*

Назовем среднее $I(f, g)$ *информацией* случайных переменных f и g.

Пример 1. Случайные переменные $f=s_1$ и $g=s_2$ *независимы* и, как следует из примера 1 пункта 4 из § 4 главы 1,

$$I(f, g)=0.$$

Пример 2. Случайные переменные $f=s_1$ и $g=s_1+s_2$ *зависимы* и, как следует из примера 2 пункта 4 из § 4 главы 1,

$$I(f, g)=1/2\cdot\log 2>0.$$

Пример 3. Случайные переменные $f=s_1-s_2$ и $g=s_1+s_2$ *зависимы.* Их значения и соответствующие значения элементарных распределений выражаются таблицами:

u	11	10	01	00
$f(u)$	0	1	−1	0
$g(u)$	2	1	1	0

;

u	11	10	01	00
$q(f(u))$	1/2	1/4	1/4	1/2
$r(g(u))$	1/4	1/2	1/2	1/4
$s(f(u),\ g(u))$	1/4	1/4	1/4	1/4

.

Информация $I(f, g)$ случайных переменных $f=s_1-s_2$ и $g=s_1+s_2$ выражается равенствами:

$$I(f, g)=\frac{1}{4}\log\frac{s(0, 2)}{q(0)\,r(2)}+\frac{1}{4}\log\frac{s(1, 1)}{q(1)\,r(1)}+$$

$$+\frac{1}{4}\log\frac{s(-1, 1)}{q(-1)\,r(1)}+\frac{1}{4}\log\frac{s(0, 0)}{q(0)\,r(0)}=$$

$$=\frac{1}{4}\left(\log\frac{1/4}{1/2\cdot 1/4}+\log\frac{1/4}{1/4\cdot 1/2}+\right.$$

$$\left.+\log\frac{1/4}{1/4\cdot 1/2}+\log\frac{1/4}{1/2\cdot 1/4}\right)=\log 2>0.$$

Замечание. В то же время коэффициент корреляции случайных переменных $f=s_1-s_2$, $g=s_1+s_2$ равен нулю и ***не выявляет** их зависимость.*

2.5. Определение

Рассмотрим конечную вероятностную модель $(U, \mathcal{A}; p, P)$ со *строго положительной* элементарной вероятностью p:

$$p(u)>0$$

для каждого исхода u из множества U.

Рассмотрим произвольные случайные переменные f и g на U. Обозначим их множество значений буквами X и Y, элементарные распределения — буквами q и r соответственно, а совместное элементарное распределение случайных переменных f и g — буквой s.

Из строгой положительности элементарной вероятности p вытекает строгая положительность элементарных распределений q и r:

$$q(x)=q(f(u))>0,\ r(y)=r(g(u))>0$$

для каждых значений $x=f(u)$ и $y=g(u)$ случайных переменных f и g. Точно так же

$$s(f(u),\ g(u))>0$$

для каждого исхода u.

При этих предположениях определена случайная переменная

$$i(f,g)=\sum_{u\in U}\log\frac{s(f(u),g(u))}{q(f(u))\,r(g(u))}\,i_u$$

на множестве исходов U со значением

$$i(f,g)(u)=\log\frac{s(f(u),g(u))}{q(f(u))\,r(g(u))}$$

для каждого исхода u.

Примеры 1—3 пунктов 1 и 4 из § 4 главы 1 поясняют определение переменной $i(f,g)$.

Случайная переменная $i(f,g)$ описывает *зависимость* случайных переменных f, g: они независимы, если и только если $i(f,g)=0$.

Оказывается, что этим свойством обладает и среднее:

$$I(f,g)=E[i(f,g)]=\sum_{u\in U}p(u)\log\frac{s(f(u),g(u))}{q(f(u))\,r(g(u))}$$

случайной переменной $i(f,g)$. Поэтому его удобно использовать в качестве *меры зависимости* случайных переменных f и g.

Определение 4. *Информацией случайных переменных f и g называется число*

$$I(f,g)=\sum_{u\in U}p(u)\log\frac{s(f(u),g(u))}{q(f(u)\,r(g(u))}.$$

Примеры 1—3 пункта 1 поясняют это определение.

При вычислении информации бывает полезно следующее предложение (суммируются слагаемые для пар x, y значений случайных переменных f, g, для которых $s(x,y)>0$):

Предложение 4. $$I(f,g)=\sum_{x,y}s(x,y)\log\frac{s(x,y)}{q(x)\,r(y)}.$$

Доказательство. Группируя слагаемые со значениями $x=f(u)$ и $y=g(u)$ случайных переменных f и g, по аналогии с доказательствами предложений 1 и 2 получаем:

$$I(f,g)=\sum_{x,y}\left(\sum_{u\in f^{-1}(x)g^{-1}(y)}p(u)\log\frac{s(x,y)}{q(x)\,r(y)}\right)=$$

$$=\sum_{x,y}\left[\left(\sum_{u\in f^{-1}(x)g^{-1}(y)}p(u)\right)\log\frac{s(x,y)}{q(x)\,r(y)}\right]=$$

$$=\sum_{x,y}s(x,y)\log\frac{s(x,y)}{q(x)\,r(y)}.$$

Предложение 4 доказано.

Пример. Рассмотрим модель Бернулли одного испытания с вероятностью успеха a $(0<a<1)$.

Для случайных переменных $f=s_1$ и $g=1-s_1$ получаем:

$$I(f, g) = s(1,0)\log\frac{s(1,0)}{q(1)r(0)} + s(0,1)\log\frac{s(0,1)}{q(0)r(1)} =$$

$$= a\log\frac{a}{a\cdot a} + (1-a)\log\frac{1-a}{(1-a)(1-a)} =$$

$$= -[a\log a + (1-a)\log(1-a)].$$

Аналогично, для случайных переменных $f=g=s_1$ получаем:

$$I(f, g) = s(1,1)\log\frac{s(1,1)}{q(1)r(1)} + s(0,0)\log\frac{s(0,0)}{q(0)r(0)} =$$

$$= -[a\log a + (1-a)\log(1-a)].$$

Как нетрудно доказать, верно следующее

Предложение. *Если случайные переменные f и g независимы, то их информация $I(f, g)$ равна нулю.*

Доказательство. По лемме 1″ пункта 5 § 3 главы 1, если случайные переменные f и g независимы, то

$$(*) \qquad s(x, y) = q(x)r(y)$$

для каждых значений $x=f(u)$ и $y=g(u)$ этих случайных переменных. И следовательно,

$$I(f, g) = \sum_u p(u)\log 1 = \sum_{x,y} s(x, y)\log 1 = 0.$$

Замечание. Из равенства $(*)$ вытекает также, что для независимых f и g случайная переменная $i(f, g)$ равна постоянной 0 и, следовательно,

$$I(f, g) = E[i(f, g)] = E(0) = 0.$$

Верно также обратное предложение: *если информация случайных переменных равна нулю, то они независимы.* Его доказательство основывается на классическом неравенстве для средних.

3.5. Теорема о средних

Среднее геометрическое каждых двух положительных чисел a и b меньше их *среднего арифметического*:

$$a^{1/2}b^{1/2} \leqslant (1/2)a + (1/2)b.$$

Причем равенство верно, если и только если $a=b$. Действительно, рассматриваемое неравенство для средних эквивалентно неравенству

$$(a^{1/2} - b^{1/2})^2 \geqslant 0.$$

Аналогичное неравенство верно для *взвешенных* средних произвольного конечного семейства положительных чисел со *строго положительными весами, в сумме равными единице.*

Рассмотрим натуральное число $n>0$, семейство $(x_j)_{1\leq j\leq n}$ положительных чисел x_j и семейство $(q_j)_{1\leq j\leq n}$ строго положительных чисел q_j, сумма которых равна единице:

$$x_j\geqslant 0,\ q_j>0,\ \Sigma q_j=1\ (1\leqslant j\leqslant n).$$

Теорема о средних. $\prod\limits_{1\leqslant j\leqslant n} x_j^{q_j}\leqslant \sum\limits_{1\leqslant j\leqslant n} q_j x_j$,

причем равенство верно, если и только если все числа x_j *равны.*

Доказательство. Теорема о средних доказывается по следующей схеме.

1. Из неравенства для среднего геометрического и принципа индукции следует, что теорема верна при

$$q_1=\ldots=q_n=1/n$$

для каждого номера $n=2^m$.

2. Из полученного результата вытекает, что теорема верна при

$$q_1=\ldots=q_n=1/n$$

для каждого номера $n>0$.

3. Из полученного результата следует, что теорема верна при каждых *рациональных* $q_1,\ldots,q_n$.

4. Из полученного результата вытекает, что неравенство теоремы верно при каждых *вещественных* $q_1,\ldots,q_n$.

5. Из результатов пунктов 3 и 4 следует, что равенство теоремы верно, если и только если $x_1=\ldots=x_n$.

Теорема оказывается полностью доказанной.

1. Случай $q_1=\ldots=q_n=1/n$ $(n=2^m)$. Рассмотрим множество всех номеров m, для которых теорема верна при $n=2^m$ и $q_1=\ldots=q_n=1/n$.

Ясно, что $0\in M$.

Для каждого номера l верно предложение: если $l\in M$, то $l+1\in M$. В самом деле, если $l\in M$, то

$$x_1^{2^{-l}}\ldots x_{2^l}^{2^{-l}}\leqslant 2^{-l}(x_1+\ldots+x_{2^l}),$$

$$x_{2^l+1}^{2^{-l}}\ldots x_{2^{l+1}}^{2^{-l}}\leqslant 2^{-l}\left(x_{2^l+1}+\ldots+x_{2^{l+1}}\right),$$

причем равенства верны тогда и только тогда, когда $x_{2^l+1}=\ldots=x_{2^{l+1}}$. Отсюда и из неравенства для среднего геометрического следует, что

$$x_1^{2^{-(l+1)}}\ldots x_{2^{l+1}}^{2^{-(l+1)}}=\left[x_1^{2^{-l}}\ldots x_{2^l}^{2^{-l}}\right]^{2^{-1}}\left[x_{2^l+1}^{2^{-l}}\ldots x_{2^{l+1}}^{2^{-l}}\right]^{2^{-1}}\leqslant$$

$$\leqslant 2^{-1}\left[2^{-l}\left(x_1+\ldots+x_{2^l}\right)+2^{-l}\left(x_{2^l+1}+\ldots+\ldots x_{2^{l+1}}\right)\right]=$$
$$=2^{-(l+1)}\left(x_1+\ldots+x_{2^{l+1}}\right).$$

Причем все равенства верны, если и только если $x_1=\ldots=x_{2^{l+1}}$. Значит, $l+1\in M$.

По принципу индукции из сказанного следует, что множество M совпадает с множеством N всех натуральных чисел. Т. е. что доказываемая теорема верна при $n=2^m$ и $q_1=\ldots=q_n=1/n$ для каждого номера $m>0$.

2. Случай $q_1=\ldots=q_n=1/n$ $(n>0)$. Для каждого натурального числа $n>0$ рассмотрим семейство $(y_k)_{1\leqslant k\leqslant 2^n}$ чисел y_k, определяемых равенствами

$$y_1=x_1,\ldots,y_n=x_n;\ y_{n+1}=\ldots=y_{2^n}=$$
$$=(1/n)(x_1+\ldots+x_n)=a.$$

Используя результат пункта 1, получаем:

$$x_1\ldots x_n a^{2^n-n}=y_1\ldots y_{2^n}\leqslant\left[2^{-n}\left(y_1+\ldots+y_{2^n}\right)\right]^{2^n}=$$
$$=(2^{-n}[x_1+\ldots+x_n+(2^n-n)a])^{2^n}=$$
$$=(2^{-n}[na+(2^n-n)a])^{2^n}=a^{2^n},$$
$$x_1\ldots x_n\leqslant a^n,$$
$$x_1^{1/n}\ldots x_n^{1/n}\leqslant a=(1/n)(x_1+\ldots+x_n).$$

Причем равенство верно, если и только если $y_1=\ldots=y_n=$ $=x_1=\ldots=x_n$. Таким образом, теорема верна при $q_1=\ldots=q_n=$ $=1/n$ для каждого номера $n>0$.

3. Случай рациональных $q_1,\ldots,q_n$. Предположим, что $q_1,\ldots,q_n$ рациональны: существуют натуральные числа $r_1,\ldots,r_n$ и натуральные числа $s_1>0,\ldots,s_n>0$ такие, что

$$q_1=r_1/s_1,\ \ldots,\ q_n=r_n/s_n.$$

Положим:

$$s=s_1,\ \ldots,\ s_n;\quad t_1=s/s_1,\ \ldots,\ t_n=s/s_n.$$

Имеем:

$$x_1^{q_1}\ldots x_n^{q_n}=x_n^{r_1/s_1}\ldots x_n^{r_n/s_n}=$$
$$=x_1^{t_1r_1/s}\ldots x_n^{t_nr_n/s}=\left(x_1^{t_1r_1}\ldots x_n^{t_nr_n}\right)^{1/s};$$
$$t_1r_1+\ldots+t_nr_n=s(q_1+\ldots+q_n)=s\cdot 1=s.$$

Рассмотрим семейство $(y_k)_{1\leqslant k\leqslant s}$ чисел y_k, определяемых равенствами

$$y_1=\ldots=y_{t_1r_1}=x_1,\ldots,y_{t_1r_1+\ldots+t_{n-1}r_{n-1}+1}=\ldots=y_s=x_n.$$

Используя результат пункта 2, получаем:

$$x_1^{q_1} \dots x_n^{q_n} = \left(x_1^{t_1 r_1} \dots x_n^{t_n r_n}\right)^{1/s} = y_1^{1/s} \dots y_s^{1/s} \leqslant$$
$$\leqslant (1/s)(y_1 + \dots + y_s) = (1/s)(t_1 r_1 x_1 + \dots + t_n r_n x_n) =$$
$$= q_1 x_1 + \dots + q_n x_n.$$

Причем все равенства верны, если и только если

$$y_1 = \dots y_s = x_1 = \dots = x_n.$$

Таким образом, теорема верна для рациональных $q_1, \dots, q_n$.

4. Случай вещественных $q_1, \dots, q_n$. Предположим, что существуют положительные числа $x_1, \dots, x_n$ и строго положительные числа $q_1, \dots, q_n$, в сумме равные 1, для которых

$$x_1^{q_1} \dots x_n^{q_n} > q_1 x_1 + \dots + q_n x_n.$$

Тогда существуют рациональные строго положительные числа $r_1, \dots, r_n$, в сумме равные 1, близкие $q_1, \dots, q_n$, для которых $x_1^{r_1} \dots x_n^{r_n}$ близко к $x_1^{q_1} \dots x_n^{q_n}$ и $r_1 x_1 + \dots + r_n x_n$ близко к $q_1 x_1 + \dots + q_n x_n$ настолько, что (рис. 34)

$$x_1^{r_1} \dots x_n^{r_n} > \frac{x_1^{q_1} \dots x_n^{q_n} + (q_1 x_1 + \dots + q_n x_n)}{2} >$$
$$> r_1 x_1 + \dots + r_n x_n$$

Рис. 34.

Это противоречит результату пункта 3. Следовательно, неравенство

$$x_1^{q_1} \dots x_n^{q_n} \leqslant q_1 x_n + \dots + q_n x_n$$

верно для всех положительных чисел $x_1, \dots, x_n$ и строго положительных чисел $q_1, \dots, q_n$, сумма которых равна 1.

Из сказанного еще не следует, что равенство верно только при равных $x_1, \dots, x_n$.

5. Равенство. Рассмотрим рациональные числа $r_1, \dots, r_n$ такие, что

$$0 < r_1 < q_1, \dots, 0 < r_n < q_n,$$

и их сумму r. Используя результаты пункта 4, получаем:

$$x_1^{q_1}\dots x_n^{q_n}=x_1^{r_1+(q_1-r_1)}\dots x_n^{r_n+(q_n-r_n)}=$$
$$=x_1^{r_1}\dots x_n^{r_n}x_1^{q_1-r_1}\dots x_n^{q_n-r_n}=$$
$$=\left[x_1^{r_1/r}\dots x_n^{r_n/r}\right]^r\left[x^{(q_1-r_1)/(1-r)}\dots x_n^{(q_n-r_n)/(1-r)}\right]^{1-r}\leqslant$$
$$\leqslant\left[(r_1/r)\,x_1+\dots+(r_n/r)\,x_n\right]^r\left[\frac{q_1-r_1}{1-r}x_1+\dots+\frac{q_n-r_n}{1-r}x_n\right]^{1-r}\leqslant$$
$$\leqslant r\left[(r_1/r)\,x_1+\dots+(r_n/r)\,x_n\right]+$$
$$+(1-r)\left[\frac{q_1-r_1}{1-r}x_1+\dots+\frac{q_n-r_n}{1-r}x_n\right]=$$
$$=r_1x_1+\dots+r_nx_n+(q_1-r_1)\,x_1+\dots+(q_n-r_n)\,x_n=$$
$$=q_1x_1+\dots+q_nx_n.$$

Из сказанного в пункте 3 следует, что все равенства верны, если и только если $x_1=\dots=x_n$: если $x_1,\dots,x_n$ различны, то первое из неравенств строгое.

Теорема о средних доказана.

Из теоремы о среднем вытекают следующие два предложения, используемые в дальнейшем.

Рассмотрим произвольные натуральное число $n>0$ и семейства $s=(s_1,\dots,s_n)$, $t=(t_1,\dots,t_n)$ строго положительных чисел $s_1,\dots,s_n$ и $t_1,\dots,t_n$, суммы $s_1+\dots+s_n$ и $t_1+\dots+t_n$ которых равны 1:

$$s_j>0,\ \Sigma s_j=1;\quad t_j>0,\ \Sigma t_j=1\ \ (1\leqslant j\leqslant n).$$

Рассмотрим функцию φ, значение которой для каждого из таких семейств t определяется равенством

$$\varphi(t)=\Sigma s_j\log t_j\ \ (1\leqslant j\leqslant n).$$

Условимся говорить, что функция φ имеет единственный *максимум в точке* s, если и только если для каждых рассматриваемых семейств t верно неравенство

$$\varphi(t)\leqslant\varphi(s)$$

и равенство

$$\varphi(t)=\varphi(s).$$

верно тогда и только тогда, когда $t=s$.

Следствие 1. *Функция φ имеет единственный максимум в точке s.*

Доказательство. Рассмотрим числа

$$x_j=t_j/s_j,\quad q_j=s_j\ \ (1\leqslant j\leqslant n).$$

Из теоремы о средних следует, что

$$\Pi (t_j/s_j)^{s_j} = \Pi x_j^{q_j} \leqslant \sum q_j x_j = \sum t_j = 1,$$
$$\sum s_j \log (t_j/s_j) \leqslant 0,$$
$$\varphi (t) = \sum s_j \log t_j \leqslant \sum s_j \log s_j = \varphi (s)$$

для каждого рассматриваемого семейства t и равенство $\varphi(t) = \varphi(s)$ верно тогда и только тогда, когда

$$t_1/s_1 = \ldots = t_n/s_n.$$

Эти равенства эквивалентны равенствам

$$t_j = c \cdot s_j \quad (1 \leqslant j \leqslant n)$$

для некоторого числа c, выражающего общее значение рассматриваемых отношений. Это число равно 1:

$$c = \Sigma s_j / \Sigma t_j = 1/1 = 1.$$

Таким образом, равенство $\varphi(t) = \varphi(s)$ верно тогда и только тогда, когда

$$t_j = s_j \quad (1 \leqslant j \leqslant n),$$

т. е. $t = s$. Следствие 1 доказано.

Рассмотрим произвольные натуральное число $n > 0$ и семейство $u = (u_1, \ldots, u_n)$ строго положительных чисел $u_1, \ldots, u_n$, сумма $u_1 + \ldots + u_n$ которых равна 1:

$$u_j > 0, \ \Sigma u_j = 1 \quad (1 \leqslant j \leqslant n).$$

Рассмотрим функцию ψ, значение которой для каждого такого семейства u определяется равенством

$$\psi(u) = -\Sigma u_j \log u_j \quad (1 \leqslant j \leqslant n).$$

Условимся говорить, что функция ψ имеет единственный *максимум в точке* $a = (1/n, \ldots, 1/n)$, если и только если для каждых рассматриваемых семейств u верно неравенство

$$\psi(u) \leqslant \psi(a)$$

и равенство

$$\psi(u) = \psi(a)$$

верно тогда и только тогда, когда $u = a$.

Следствие 2. *Функция ψ имеет единственный максимум в точке a.*

Доказательство. Рассмотрим числа

$$s_j = u_j, \quad t_j = 1/n \quad (1 \leqslant j \leqslant n)$$

и семейства $s = u$, $t = a$.

Из **следствия** 1 теоремы о средних вытекает, что

$$\psi(u)=-\Sigma u_j \log u_j=-\Sigma s_j \log s_j \leqslant$$
$$\leqslant -\Sigma s_j \log t_j = -\Sigma u_j \log 1/n =$$
$$=-(\Sigma u_j) \log 1/n = -1 \cdot \log 1/n = -(\Sigma 1/n) \log 1/n =$$
$$=-\Sigma 1/n \log 1/n = \psi(a)$$

и равенство $\psi(u)=\psi(a)$ верно тогда и только тогда, когда $u=a$.

Замечание. $\psi(a)=-\Sigma 1/n \log 1/n=\log n$.

4.5. Теорема об информации

Рассмотрим конечную вероятностную модель $(U, \mathscr{A}; p, P)$ и случайные переменные f, g, описанные в пункте 2 при определении информации $I(f, g)$. Основные свойства информации выражает

Теорема об информации. *Информация случайных переменных f и g положительна. Она равна нулю, если и только если случайные переменные f и g независимы.*

Доказательство. Рассмотрим семейства s и t чисел $s(x, y)$ и $t(x, y)=q(x)\,r(y)$, полученные произвольной нумерацией пар значений случайных переменных f, g. Из предложения 4 следует, что

$$I(f, g)=\Sigma s(x, y) \log s(x, y) -$$
$$-\Sigma s(x, y) \log t(x, y).$$

По следствию 1 теоремы о средних отсюда вытекает, что

$$I(f, g) \geqslant 0$$

и равенство

$$I(f, g)=0$$

верно, если и только если

$$s(x, y)=t(x, y)=q(x)\,r(y)$$

для всех пар x, y значений случайных переменных f, g. По лемме 1″ пункта 5 § 3 главы 1, эти равенства эквивалентны независимости случайных переменных f, g. Теорема доказана.

Непосредственно из теоремы об информации вытекает следующий

Критерий зависимости. *Случайные переменные f и g зависимы, если и только если их информация строго положительна.*

Примеры 1—3 пункта 1 поясняют этот критерий.

§ 6*. ЭНТРОПИЯ

Как и термин *информация,* термин *энтропия* используется здесь в очень узком и точно определенном смысле. Энтропия случайной переменной является важной характеристикой ее распределения.

Она служит *мерой неопределенности* этого распределения: в случае вырожденного распределения энтропия минимальна, а в случае равномерного — максимальна.

Энтропия определяется как информация равных случайных переменных.

1.6. Примеры

Рассмотрим несколько примеров, поясняющих используемое понятие энтропии. Возьмем примеры пункта 1 § 5 для модели Бернулли $n=2$ испытаний с вероятностью успеха $a=1/2$.

Если $f=g$, то

$$q(f(u))=r(g(u))=s(f(u),\ g(u))$$

для каждого исхода u из множества $U=\{11, 10, 01, 00\}$.
Поэтому

$$h(f)=i(f, f)=-[\log q(f(11))\, i_{11}+$$
$$+\log q(f(10))\, i_{10}+\log q(f(01))\, i_{01}+$$
$$\log q(f(00))\, i_{00}],$$
$$H(f)=E[h(f)]=I(f, f)=[p(11)\log q(f(11))+$$
$$+p(10)\log q(f(10))+p(01)\log q(f(01))+$$
$$+p(00)\log q(f(00))].$$

Назовем среднее $H(f)$ энтропией случайной переменной f.

Пример 1.

$$H(s_1)=H(s_2)=-1/4\ [2\log q(1)+2\log q(0)]=$$
$$=[1/2\log 1/2+1/2\log 1/2]=\log 2.$$

Пример 2.

$$H(s_1+s_2)=-1/4[\log q(2)+2\log q(1)+\log q(0)]=$$
$$=[1/4\log 1/4+1/2\log 1/2+1/4\log 1/4]=$$
$$=3/2\log 2.$$

Пример 3.

$$H(s_1-s_2)=-1/4[2\log q(0)+\log q(1)+\log q(-1)]=$$
$$=[1/2\log 1/2+1/4\log 1/4+1/4\log 1/4]=$$
$$=3/2\log 2.$$

Пример 4. Рассмотрим постоянную $f=1$, имеющую вырожденное распределение q со значением $q(1)=1$.
Имеем:

$$H(1)=-1/4[4\log q(1)]=-1\cdot\log 1=0.$$

Энтропия постоянной равна нулю. Постоянная не имеет неопределенности.

2.6. Определение

Рассмотрим конечную вероятностную модель $(U, \mathcal{A}; p, P)$ со *строго положительной* элементарной вероятностью p и произвольную случайную переменную f на U с множеством значений X и элементарным распределением q.

Для случайной f определена случайная переменная

$$h(f) = i(f, f) = -\sum_{u \in U} \log q(f(u))\, i_u$$

на множестве исходов U со значением

$$h(f(u)) = -\log q(f(u))$$

для каждого исхода u.

Примеры 1—4 пункта 1 поясняют определение переменной $h(f)$.

Среднее $H(f)$ случайной переменной $h(f)$ выражается равенствами

$$H(f) = E(h(f)) = I(f, f) = -\sum_{u \in U} p(u) \log q(f(u)).$$

Сформулируем

Определение 5. *Энтропией случайной переменной называется число*

$$H(f) = I(f, f).$$

Примеры 1—4 пункта 1 поясняют это определение.

При вычислении энтропии бывает полезно

Предложение 5. $H(f) = -\sum_{x} q(x) \log q(x)$.

Это равенство вытекает из равенства предложения 4 и равенств

$$s(x, x) = r(x) = q(x)$$

для каждого значения x случайной переменной f.

Замечание. Энтропия случайной переменной положительна. Это следует из положительности информации и предложения 5.

Пример. Как было показано в пункте 2 § 5 для модели Бернулли $n=2$ испытаний с вероятностью успеха $a (0 < a < 1)$, верно равенство

$$H(s_1) = -[a \log a + (1-a) \log (1-a)].$$

3.6. Теорема об энтропии

Рассмотрим конечную вероятностную модель $(U, \mathcal{A}; p, P)$ и случайные переменные f и g, описанные в пункте 2 § 5 при определении информации. Основные свойства энтропии выражает

Теорема об энтропии. *Информация случайных переменных f и g меньше энтропии каждой из них. Энтропия произвольной случайной переменной меньше логарифма числа ее значений и*

равна ему, если и только если переменная имеет равномерное распределение.

Доказательство. 1. Заметим, что

$$q(x)=\sum_y s(x, y), r(y)=\sum_x s(x, y), r(y)\geqslant s(x, y)$$

и, значит,

$$\frac{s(x, y)}{q(x)\, r(y)}\leqslant\frac{1}{q(x)}$$

для каждых значений x, y случайных переменных f, g. Отсюда и из предложений 4 и 5 следует, что

$$I(f, g)=\sum_{x,y} s(x, y)\log\frac{s(x, y)}{q(x)\, r(y)}\leqslant$$

$$\leqslant\sum_{x,y} s(x, y)\log\frac{1}{q(x)}=-\sum_x\Bigl(\sum_y s(x, y)\Bigr)\log q(x)=$$

$$=-\sum_x q(x)\log q(x)=H(f).$$

Аналогично доказывается, что

$$I(f, g)\leqslant H(g).$$

Таким образом,

$$I(f, g)\leqslant\min\{H(f), H(g)\}.$$

2. Второе утверждение теоремы об энтропии вытекает из предложения 5 и следствия 2 теоремы о средних. Обозначив буквой n число значений случайной переменной f, получим:

$$H(f)=-\sum_x q(x)\log q(x)\leqslant$$

$$\leqslant\sum_x 1/n\log 1/n=\log n,$$

причем равенство

$$H(f)=\log n$$

верно, если и только если

$$q(x)=1/n$$

для каждого значения x случайной переменной f.

4.6. Примеры

Рассмотрим еще несколько примеров, поясняющих понятие информации и энтропии.

Условимся использовать логарифмы с основанием 2 и соответствующую единицу информации называть битом.

Пример 1. Заблудившаяся корова. В одном из 64 квадратов поля спит заблудившаяся корова. Ее ищет пастух, который из $k_1=64$ квадратов поля выбирает l_1 квадратов и осматривает их. Если корова находится, то поиски на этом заканчиваются. Если нет, то из оставшихся $k_2=k_1-l_1$ квадратов поля пастух выбирает l_2 квадратов и осматривает их. И так далее, пока корова не находится.

Перебрав все возможные планы поисков, можно убедиться в том, что план последовательного деления пополам ($l_1=32, l_2=16, l_3=8, l_4=4, l_5=2, l_6=1$) минимизирует наибольшее число осматриваемых квадратов, нужное для того, чтобы найти корову в любом квадрате поля.

План последовательного деления пополам является оптимальным также в том смысле, что он при каждом осмотре выбранных квадратов поля обеспечивает максимум информации о корове, т. е. соответствует наибольшей зависимости между действиями пастуха и местонахождением коровы.

Рассмотрим стандартную модель Лапласа $k=64$ равномерных исходов $1, \ldots, k$. Будем считать, что случайная переменная f со значением $f(u)=u$ для каждого исхода u описывает номер квадрата поля, в котором находится корова. Рассмотрим также случайную переменную g, равную индикатору множества L номеров квадратов поля, выбираемых пастухом для осмотра из данных k квадратов. Число выбираемых квадратов обозначим буквой l.

Как нетрудно проверить,

$$X=\{1, \ldots, k\}, \quad q(1)=\ldots=q(k)=1/k;$$

$$Y=\{1, 0\}, \ r(1)=l/k, \ r(0)=1-l/k;$$

$$s(x, 1)=1/k \, (x \in L), \quad s(x, 1)=0 \ (x \notin L),$$

$$s(x, 0)=0 \ (x \in L), \ s(x, 0)=1/k \ (x \notin L).$$

Используя равенства предложений 4 и 5, получаем:

$$I(f,g) = l/k \cdot \log \frac{1/k}{1/k \cdot l/k} + \frac{k-l}{k} \log \frac{1/k}{1/k\,(1-l/k)} =$$

$$= -[l/k \cdot \log (l/k) + (1-l/k) \log (1-l/k] = H(g).$$

Из теоремы об энтропии следует, что максимум информации достигается, если и только если

$$l/k=1/2, \ l=k/2.$$

В этом случае

$$I(f, \ g)=H(g)-\log 2.$$

Таким образом, план последовательного деления пополам при каждом осмотре квадратов позволяет получить 1 бит информации о корове. Шесть таких осмотров полностью устраняют неопределенность:

$$H(f)=\log 64=6 \text{ битов}.$$

нахождения животного. Любой другой план хотя бы на одном шаге дает меньшую информацию.

Пример 2. Отгадывание числа. Из $k_1=2^n$ чисел выбирается одно. Для определения этого числа отгадывающий применяет следующую процедуру. Из данных k_1 чисел он выбирает l_1 чисел. Задумавший число сообщает, находится оно среди выбранных или нет. Тем самым определяются k_2 чисел, среди которых находится задуманное ($k_2=k_1$ либо $k_2=k_1-l_1$). Из этих k_2 чисел отгадывающий выбирает l_2 чисел. Задумавший число снова сообщает, находится оно среди выбранных или нет. И так далее, пока задуманное число не окажется единственным выбранным числом.

Рассуждения, аналогичные проведенным в примере 1, показывают, что оптимальной процедурой при отгадывании задуманного числа является процедура деления множества чисел, среди которых находится задуманное число, пополам. На какие именно равные половины делить эти числа, не существенно. Такая процедура последовательного деления пополам обеспечивает получение на каждом шаге максимума информации о задуманном числе.

При последовательном делении пополам отгадывающий каждый раз получает 1 бит информации о задуманном числе. При других процедурах информация меньше. Например, если из $k_1=2^4=16$ выбирать $l=1$ число, то получаемая при этом информация меньше 0,3 бита:

$$I(f,\ g)=-(1/16\cdot\log 1/16+15/16\cdot\log 15/16)<0{,}3.$$

Замечание. Этот пример подчеркивает вероятностный характер данного определения информации. При реальном таком отгадывании можно случайно сразу выбрать задуманное число.

Пример 3. Линия связи. Простейшие источник сигналов и линия их передачи (например, один такт работы электрического телеграфа) можно описать следующей вероятностной моделью.

Рассмотрим множество исходов $U=\{00,\ 01,\ 10,\ 11\}$. Исход $u=xy$ описывает передачу сигнала x и прием сигнала y. Событие $A_x=\{x0,\ x1\}$ описывает передачу сигнала x, а событие $B_y=\{0y,\ 1y\}$ — прием сигнала y ($x=0,\ 1;\ y=0,\ 1$).

Элементарную вероятность p определим равенствами

$$p(x,\ y)=a_x t_{xy}\ (x=0,\ 1;\ y=0,\ 1),$$

в которых a_x и t_{xy} обозначают числа такие, что

$$a_0>0,\ a_1>0,\ a_0+a_1=1;$$

$$t_{00}>0,\ \ t_{01}>0,\ \ t_{00}+t_{01}=1;$$

$$t_{10}>0,\ t_{11}>0,\ t_{10}+t_{11}=1.$$

Из этого определения следует, что для каждых значений $x=0,\ 1$ и $y=0,\ 1$ верны равенства:

$$P(A_x)=p(x0)+p(x1)=a_x t_{x0}+a_x t_{x1}=a_x(t_{x0}+t_{x1})=a_x,$$

$$P_{A_x}(B_y)=\frac{P(A_x B_x)}{P(A_x)}=\frac{a_x t_{xy}}{a_x}=t_{xy},$$

$$P(B_y) = P(A_0) P_{A_0}(B_y) + P(A_1) P_{A_1}(B_y) = a_0 t_{0y} + a_1 t_{1y}.$$

Рассмотрим случайные переменные f и g со значениями

$$f(xy) = x,\ g(xy) = y\ \ (x=0,\ 1;\ y=0,\ 1).$$

Эти переменные описывают соответственно переданный и принятый сигналы. Имеем:

$$X = \{0,\ 1\},\ q(x) = a_x\ \ (x=0,\ 1);$$

$$Y = \{0,\ 1\},\ r(y) = a_0 t_{0y} + a_1 t_{1y}\ \ (y=0,\ 1),$$

$$s(x,\ y) = a_x t_{xy}\ \ (x=0,\ 1;\ y=0,\ 1).$$

Используя предложение 4, получаем:

$$I(f,\ g) = \sum_{x,y} a_x t_{xy} \log \frac{t_{xy}}{a_0 t_{0y} + a_1 t_{1y}}.$$

В частности, если

$$t_{00} = t_{01} = t_{10} = t_{11} = 1/2,$$

то переданный и принятый сигналы независимы и

$$I(f,\ g) = 0.$$

Глава 3

ЗАКОН БОЛЬШИХ ЧИСЕЛ

Среднее случайной переменной тесно связано со средним арифметическим ее реально наблюдаемых значений. Если наблюдение повторяется достаточно много раз, то, как правило, среднее арифметическое реально наблюдаемых значений случайной переменной оказывается близким к ее среднему. В частности, реально наблюдаемая частота появления события оказывается близкой к его вероятности. Например, частота появления герба при подбрасывании симметричной монеты, как правило, оказывается близкой к 1/2. Соответствующее формальное предложение называется законом больших чисел.

Этот закон утверждает, что вероятность сколько-нибудь значительного отклонения *арифметического* среднего достаточно *большого числа* независимых и одинаково распределенных случайных переменных от их *вероятностного* среднего произвольно *мала*. Закон больших чисел позволяет оценить вероятностное среднее с помощью арифметического. В частности, оценить *вероятность* с помощью *частоты*.

Различные количественные выражения для закона больших чисел получаются из классического *неравенства Чебышева*, ко-

торое оценивает вероятность значительного отклонения случайной переменной от ее среднего.

При первом чтении рекомендуется ограничиться параграфами 1—4.

Задачи к главе 3 собраны в § 2 главы 2 части III.

§ 1. ПРИМЕРЫ

Рассмотрим несколько примеров, поясняющих содержание закона больших чисел.

1.1. Пример 2. Подбрасывание монеты

Рассмотрим опыт, состоящий из последовательности n испытаний, каждое из которых заключается в подбрасывании симметричной монеты один раз и имеет два возможных результата: появление герба либо появление цифры. Такой опыт описывается *моделью Бернулли n испытаний с вероятностью успеха $a=1/2$*. Для каждого номера $j=1,\dots, n$ случайная переменная $s_j = i_{\mathrm{y}_j}$ описывает число гербов при j-м подбрасывании монеты, арифметическое среднее $(1/n)\,s=(1/n)\sum\limits_{1\leqslant j\leqslant n} s_j$ этих случайных переменных — частоту герба при n подбрасываниях. Вероятностное среднее каждой из независимых и одинаково распределенных случайных переменных s_j равно вероятности $a=1/2$.

Закон больших чисел в этом примере выражается в том, что при достаточно *большом числе* подбрасываний вероятность сколько-нибудь значительного отклонения частоты $(1/n)s$ от $a=1/2$ произвольно *мала*. Это значит, что при большом числе подбрасываний наблюдаемая *частота* герба будет, как правило, *близка вероятности* 1/2.

Один из основоположников математической статистики К. Пирсон подбросил монету $n=24\,000$ раз. Герб появился 12 012, а цифра — 11 988 раз. Частота герба оказалась равной $12\,012/24\,000=0{,}5005$. Она отличается от вероятности 1/2 на 0,0005.

Находясь в лагере для военнопленных, Д. Керрих провел десять опытов с монетой по $n=1\,000$ подбрасываний каждый. Результаты этих опытов содержит таблица

$(1/n)s$	0,502	0,511	0,497	0,529	0,504	0,476	0,507	0,528	0,504	0,529
$\lvert(1/n)s-a\rvert$	0,002	0,011	0,003	0,029	0,004	0,024	0,007	0,028	0,004	0,029

2.1. Пример 2. Случайные числа

Рассмотрим опыт, состоящий из последовательности n испытаний, каждое из которых заключается в выборе наугад одной из цифр 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.

Такой опыт удобно описать *моделью Лапласа*, множеством исходов для которой служит множество всех строк длины n, составленных из цифр (n-значных чисел). Для каждого номера $j=1, \ldots, n$ рассмотрим случайную переменную f_{ij}, значение $f_{ij}(u)$ которой равно 1 для каждой строки u с цифрой i на j-м месте и равно 0 для всех остальных исходов. Случайная переменная f_{ij} описывает число цифр i при j-м выборе, а арифметическое среднее $\bar{f}_i = (1/n) \sum_{1 \leqslant j \leqslant n} f_{ij}$ этих случайных переменных — частоту цифры i при n выборах. Для каждой цифры i вероятностное среднее каждой из независимых и одинаково распределенных случайных переменных f_{ij} равно вероятности $a_i = 1/10$ выбрать из цифр 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 наугад цифру i.

Закон больших чисел в этом примере выражается в том, что для каждой цифры i при достаточно *большом числе* выборов вероятность сколько-нибудь значительного отклонения частоты $\bar{f}_i$ от $a_i = 1/10$ произвольно *мала*. Это значит, что при большом числе выборов наблюдаемая *частота* данной цифры будет, как правило, *близка вероятности* 1/10.

По опубликованным таблицам, содержащим $n = 1\,000\,000$ выбранных наугад цифр, их частоты оказались следующими:

i	0	1	2	3	4
$\bar{f}_i$ ···	0,099802	0,100050	0,100641	0,100311	0,100094
$\bar{f}_i - a_i$	−0,000198	0,000050	0,000641	0,000311	0,000094
i	5	6	7	8	9
$\bar{f}_i$ ···	0,10214	0,099942	0,099559	0,100107	0,096280
$\bar{f}_i$ ···	0,000214	−0,000058	0,000441	0,000107	−0,003720

3.1. Пример 3. Измерения

Рассмотрим опыт, состоящий из последовательности n измерений некоторой величины. Предположим, что эти измерения с учетом ошибок, зависящих от различных причин, можно описать *независимыми* и *одинаково распределенными* случайными переменными $f_j (1 \leqslant j \leqslant n)$ на множестве исходов U конечной вероятностной модели (U, $\mathscr{A}$; p, P) и что среднее a каждой из случайных переменных f_j равно значению измеряемой величины.

Закон больших чисел в этом примере выражается в том, что при достаточно *большом числе* измерений вероятность сколько-нибудь значительного отклонения среднего арифметического $\bar{f} = (1/n) \sum_{1 \leqslant j \leqslant n} f_j$ от a произвольно *мала*. Это значит, что при большом числе измерений получаемое *среднее арифметическое* результатов будет, как правило, *близко значению* измеряемой величины.

В опыте Милликена по измерению заряда электрона $n=58$ измерений дали следующие результаты (в 10^{-10} абс. эл.-ст. ед.):

4,781	4,764	4,777	4,809	4,761	4,769
4,795	4,776	4,765	4,790	4,792	4,806
4,769	4,771	4,785	4,779	4,758	4,779
4,792	4,789	4,805	4,788	4,764	4,785
4,779	4,772	4,768	4,772	4,810	4,790
4,775	4,789	4,801	4,791	4,799	4,777
4,772	4,764	4,785	4,788	4,799	4,749
4,791	4,774	4,783	4,783	4,797	4,781
4,782	4,778	4,808	4,740	4,790	
4,767	4,791	4,771	4,775	4,747	

Предположив, что для опыта Милликена верна рассматриваемая модель, и вычислив среднее арифметическое приведенных в таблице результатов, можно получить приблизительное значение заряда электрона: $4{,}781 \cdot 10^{-10}$ абс. эл.-ст. ед.

§ 2. НЕРАВЕНСТВО ЧЕБЫШЕВА

Это неравенство оценивает вероятность значительного отклонения случайной переменной от ее среднего. Оценка производится с помощью дисперсии.

1.2. Пример

Прежде, чем сформулировать и доказать неравенство Чебышева, рассмотрим поясняющий пример.

Среднее $E(f)$ частоты $f=(1/3)s$ успеха в модели Бернулли $n=3$ испытаний с вероятностью успеха $a=1/2$ равно вероятности a. Дисперсия $D(f)$ равна $(1/3)a(1-a)=1/12$. Рассмотрим число $\varepsilon=1/3$ и условимся считать, что *значительное* отклонение случайной переменной f от ее среднего $E(f)$ описывается событием

$$B=B(f,\ \varepsilon)=\{u: |f(u)-E(f)|\geqslant\varepsilon\}=\{u: |(1/3)s(u)-1/2|\geqslant 1/3\}=$$
$$=\{111,\ 000\}.$$

Оценим вероятность $P(B)$ этого события с помощью дисперсии $D(f)$.

По определению, дисперсия $D(f)$ равна среднему квадратичного отклонения $g=(f-E(f))^2$ случайной переменной f:

$$D(f)=E(g)=E[(f-E(f))^2]=\Sigma(f(u)-E(f))^2p(u)=$$
$$=(f(111)-1/2)^2 1/8+(f(110)-1/2)^2 1/8+(f(101)-1/2)^2 1/8+$$
$$+(f(011)-1/2)^2 1/8+(f(100)-1/2)^2 1/8+(f(010)-1/2)^2 1/8+$$
$$+(f(001)-1/2)^2 1/8+(f(000)-1/2)^2 1/8=(1-1/2)^2 1/8+$$
$$+3(2/3-1/2)^2 1/8+3(1/3-1/2)^2 1/8+(0-1/2)^2 1/8=1/12.$$

Полное квадратичное отклонение g случайной переменной f больше, чем частичное $h=gi_B=gi_{\{111,000\}}$:

$$g(u)=h(u) \quad (u \in B),\ g(u) \geqslant 0=h(u) \quad (u \in B'),$$

т. е.

$$g(111)=h(111)=(1-1/2)^2=1/4,\ g(000)=h(000)=(0-1/2)^2= \\ =1/4,\ g(110)=g(101)=g(011)=(2/3-1/2)^2=1/36 \geqslant 0= \\ =h(101)=h(011),\ g(100)=g(010)=g(001)=(1/3-1/2)^2= \\ =1/36 \geqslant 0=h(100)=h(010)=h(001).$$

В то же время, по определению события B, описывающего значительные отклонения,

$$h \geqslant \varepsilon^2 i_B=(1/9)\, i_{\{111,000\}},$$
$$h(111)=h(000)=1/4 \geqslant 1/9,$$
$$h(110)=h(101)=h(011)=h(100)=h(010)=h(001)=0.$$

По правилу неравенства для среднего из неравенств $g \geqslant gi_B \geqslant \geqslant \varepsilon^2 i_B$ следует, что

$$D(f)=E(g) \geqslant E(gi_B)=(1-1/2)^2 1/8+(0-1/2)^2 1/8 \geqslant \\ \geqslant E(\varepsilon^2 i_B)=(1/9)\, 1/8+(1/9)\, 1/8.$$

Правило умножения на постоянную и правило индикатора позволяют иначе вычислить последнее среднее:

$$E(\varepsilon^2 i_B)=\varepsilon^2 E(i_B)=\varepsilon^2 P(B)=(1/9)\,(1/8+1/8)=(1/9)\,(1/4).$$

Таким образом,

$$1/12=D(f) \geqslant \varepsilon^2 P(B)=(1/9)\,(1/4)$$

и, следовательно,

$$1/4=P(B(f,\ \varepsilon)) \leqslant \varepsilon^{-2} D(f)=9(1/12)=3/4.$$

Получена оценка вероятности события $B=B(f,\ \varepsilon)$ с помощью дисперсии $D(f)$.

Событие $A=A(f,\ \varepsilon)$, дополнительное к $B=B(f,\ \varepsilon)$, описывает *незначительное* отклонение случайной переменной f от ее среднего $E(f)$:

$$A=A(f,\ \varepsilon)=\{u: |f(u)-E(f)|<\varepsilon\}=\{u: |(1/3)s(u)-1/2|<1/3\}= \\ =\{110,\ 101,\ 011,\ 100,\ 010,\ 001\}.$$

По правилу дополнения, из полученного неравенства для вероятности события B следует, что

$$3/4=P(A(f,\ \varepsilon)) \geqslant 1-P(B(f,\ \varepsilon)) \geqslant 1-(1/\varepsilon^2)D(f)=1-9(1/12)=1/4.$$

Замечание. Из этого примера видно также, что полученные оценки довольно грубы.

2.2. Доказательство неравенства Чебышева

Рассмотрим вероятностную модель $(U, \mathcal{A}; p, P)$ и случайную переменную f на множестве исходов U, имеющую среднее $E(f)$ и дисперсию $D(f)$. Для каждого числа $\varepsilon>0$ верно

Неравенство Чебышева:

$$P\{u: |f(u)-E(f)| \geqslant \varepsilon\} \leqslant \varepsilon^{-2} D(f).$$

Доказательство. Рассмотрим событие

$$B(f, \varepsilon)=\{u: |f(u)-E(f)| \geqslant \varepsilon\},$$

составленное из всех исходов u, для которых значение $f(u)$ случайной переменной f отличается от ее среднего $E(f)$ по абсолютной величине больше чем на ε. Для каждого исхода u из множества U верны неравенства

$$(f(u)-E(f))^2 \geqslant (f(u)-E(f))^2 i_{B(f, \varepsilon)}(u) \geqslant \varepsilon^2 i_{B(f, \varepsilon)}(u),$$

т. е. неравенства

$$(f-E(f))^2 \geqslant (f-E(f))^2 i_{B(f, \varepsilon)} \geqslant \varepsilon^2 i_{B(f, \varepsilon)}.$$

По определению дисперсии $D(f)$ и правилу неравенства для среднего, из этих неравенств следует, что

$$D(f)=E(f-E(f))^2 \geqslant E(\varepsilon^2 i_{B(f, \varepsilon)}).$$

Из правила умножения на число и правила индикатора вытекает, что

$$E(\varepsilon^2 i_{B(f, \varepsilon)})=\varepsilon^2 E(i_{B(f, \varepsilon)})=\varepsilon^2 P(B(f, \varepsilon)).$$

Таким образом,

$$D(f) \geqslant \varepsilon^2 P(B(f, \varepsilon))$$

и, значит,

$$P(B(f, \varepsilon)) \leqslant \varepsilon^{-2} D(f),$$

что и требовалось доказать.

3.2. Следствие

Из доказанного неравенства и правила дополнения для вероятности вытекает эквивалентное неравенство, которое тоже называется *неравенством Чебышева.*

Следствие.

$$P\{u: |f(u)-E(f)| < \varepsilon\} \geqslant 1-\varepsilon^{-2} D(f).$$

Доказательство. Рассмотрим событие

$$A(f, \varepsilon)=U-B(f, \varepsilon)=\{u: |f(u)-E(f)| < \varepsilon\},$$

составленное из всех исходов u, для которых значение $f(u)$ случайной переменной f отличается от ее среднего $E(f)$ по абсолют-

ной величине строго меньше чем на ε. Используя доказанное неравенство и правило дополнения, получаем:

$$P(A(f, \varepsilon)) = 1 - P(B(f, \varepsilon)) \geqslant 1 - \varepsilon^2 D(f),$$

что и требовалось доказать.

Пример пункта 1.2 поясняет сами полученные неравенства и их доказательства.

Замечание. При слишком малых числах ε полученные неравенства могут быть тривиальными: правая часть первого может быть больше 1, а второго — меньше 0. Так будет, в частности, если в примере пункта 1.2 взять $\varepsilon = 1/4$. Хотя в этом случае попрежнему $P(B(f, \varepsilon)) = 1/4$ и $P(A(f, \varepsilon)) = 3/4$.

§ 3. ТЕОРЕМА ЧЕБЫШЕВА

Эта теорема следует из неравенства Чебышева и оценивает вероятность значительного отклонения *арифметического среднего* независимых и одинаково распределенных случайных переменных от их *вероятностного среднего.*

1.3. Доказательство теоремы

Рассмотрим конечную вероятностную модель $(U, \mathscr{A}; p, P)$ и *попарно независимые* и *одинаково распределенные* случайные переменные $f_j (1 \leqslant j \leqslant n)$ на множестве исходов U. Обозначим среднее каждой из случайных переменных f_j буквой a, а дисперсию — буквой b^2:

$$E(f_j) = a, \quad D(f_j) = b^2 \quad (1 \leqslant j \leqslant n).$$

Рассмотрим также среднее арифметическое

$$f = n^{-1} \sum_{1 \leqslant j \leqslant n} f_j$$

случайных переменных f_j. Вычислим среднее и дисперсию случайной переменной f.

Лемма. $E(f) = a$, $D(f) = n^{-1} b^2$.

Доказательство. Используя общее правило линейности для среднего и одинаковую распределенность случайных переменных f_j, получаем:

$$E(f) = E\left(n^{-1} \sum_{1 \leqslant j \leqslant n} f_j\right) = n^{-1} \sum_{1 \leqslant j \leqslant n} E(f_j) = n^{-1} na = a.$$

2. Используя правило умножения на число для дисперсии, общее правило для дисперсии суммы семейства попарно независимых случайных переменных и одинаковую распределенность случайных переменных f_j, получаем:

$$D(f) = D\left(n^{-1} \sum_{1 \leqslant j \leqslant n} f_j\right) = n^{-2} \sum_{1 \leqslant j \leqslant n} (Df_j) = n^{-2} n b^2 = n^{-1} b^2.$$

Лемма доказана.

При сделанных предположениях и введенных обозначениях из этой леммы и неравенства Чебышева вытекает, что для каждого числа $\varepsilon > 0$ верна

Теорема Чебышева.

$$P\left\{u: \left| n^{-1} \sum_{1 \leqslant j \leqslant n} f_j(u) - a \right| \geqslant \varepsilon \right\} \leqslant \varepsilon^{-2} n^{-1} b^2.$$

Доказательство. Рассмотрим случайную переменную

$$f = n^{-1} \sum_{1 \leqslant j \leqslant n} f_j.$$

По лемме, ее среднее $E(f)$ равно числу a. Следовательно, событие, вероятность которого нужно оценить, описывает значительное отклонение случайной переменной f от ее среднего $E(f)$:

$$B(f, \varepsilon) = \left\{u: \left| n^{-1} \sum_{1 \leqslant j \leqslant n} f_j(u) - a \right| \geqslant \varepsilon \right\} = \{u: |f(u) - E(f)| \geqslant \varepsilon\}.$$

Поэтому для оценки вероятности этого события можно использовать неравенство Чебышева. Вместе с равенством леммы для дисперсии $D(f)$ оно показывает, что

$$P(B(f, \varepsilon)) \leqslant \varepsilon^{-2} D(f) = \varepsilon^{-2} n^{-1} b^2.$$

Теорема доказана.

Замечание. Для каждого $\varepsilon > 0$ при всех достаточно *больших числах* n правая часть $\varepsilon^{-2} n^{-1} b^2$ полученного неравенства произвольно *мала*. Вместе с ней произвольно мала и оцениваемая ею вероятность отклонения *арифметического* среднего от *вероятностного* больше чем на ε, т. е. верен *закон больших чисел*.

2.3. Следствие

Из доказанного неравенства и правила дополнения для вероятности вытекает эквивалентное неравенство. Условимся его тоже называть *теоремой Чебышева*.

Следствие.

$$P\left\{u: \left| n^{-1} \sum_{1 \leqslant j < n} f_j(u) - a \right| < \varepsilon \right\} \geqslant 1 - \varepsilon^{-2} n^{-1} b^2.$$

Это неравенство оценивает вероятность *незначительного* отклонения арифметического среднего от вероятностного.

Замечание. Для каждого $\varepsilon > 0$ при всех достаточно *больших числах* n правая часть $1 - \varepsilon^{-2} n^{-1} b^2$ полученного неравенства произвольно *близка* к 1. Вместе с ней произвольно близка к 1 и оцениваемая ею вероятность отклонения *арифметического* среднего от *вероятностного* строго меньше чем на ε. Это другая форма *закона больших чисел*.

3.3. Пример. Точность и число измерений

Как и в § 1, рассмотрим опыт, состоящий из последовательности n измерений некоторой величины. По-прежнему будем предполагать, что эти измерения можно описать независимыми и одинаково распределенными случайными переменными f_j на множестве исходов U конечного вероятностного пространства $(U, \mathscr{A}; p, P)$ и что среднее a каждой из случайных переменных f_j равно значению измеряемой величины.

Дисперсия b случайных переменных f_j характеризует *точность* измерений.

На основании закона больших чисел можно считать, что при сделанных предположениях a приближенно равно арифметическому среднему результатов измерений.

Теорема Чебышева позволяет оценить точность ε такого приближения.

Рассмотрим число $\alpha>0$ и условимся считать события, вероятность которых в используемой модели меньше α, практически невозможными, а события, вероятность которых больше $1-\alpha$,— практически достоверными.
Неравенство

$$\varepsilon^{-2}n^{-1}b^2\leqslant\alpha$$

верно, если

$$\varepsilon=(n\alpha)^{-1/2}b.$$

По теореме Чебышева отсюда следует, что при такой точности ε отклонение арифметического среднего рассматриваемых случайных переменных от вероятностного больше чем на ε, практически невозможно. Эта точность ε практически достоверна.

В частности, если $b^2=1$, $n=10^3$ и $\alpha=10^{-3}$, то $\varepsilon=1$.

Теорема Чебышева позволяет также оценить число измерений n, практически достоверно обеспечивающее нужную точность. Из полученного выше неравенства следует, что

$$n\geqslant\varepsilon^{-2}\alpha^{-1}b^2.$$

При таком числе измерений n отклонение арифметического среднего рассматриваемых случайных переменных от вероятностного больше чем на ε практически невозможно.

В частности, если $b^2=10^{-2}$, $\alpha=10^{-2}$ и $\varepsilon=10^{-1}$, то $n\geqslant10^2$.

§ 4. ТЕОРЕМА БЕРНУЛЛИ

Эта теорема является следствием теоремы Чебышева для модели Бернулли. Она оценивает вероятность значительного отклонения *частоты* успеха от его *вероятности*.

1.4. Доказательство теоремы Бернулли

Рассмотрим модель Бернулли n испытаний с вероятностью успеха a и случайную переменную $n^{-1}s = n^{-1} \sum_{1 \leqslant j \leqslant n} s_j$, равную арифметическому среднему независимых и одинаково распределенных случайных переменных s_j, описывающих *число* успехов при j-м испытании $(1 \leqslant j \leqslant n)$. Случайная переменная $n^{-1}s$ описывает *частоту* успеха при n испытаниях. Для каждого номера $j = 1, \ldots, n$ верны равенства

$$E(s_j) = a, \; D(s_j) = a(1-a).$$

Отсюда и из теоремы Чебышева вытекает, что для каждого $\varepsilon > 0$ верна

Теорема Бернулли.

$$P\left\{u : \left| n^{-1} \sum_{1 \leqslant j \leqslant n} s_j(u) - a \right| \geqslant \varepsilon \right\} \leqslant \varepsilon^{-2} n^{-1} a(1-a).$$

Замечание. Для каждого $\varepsilon > 0$ при всех достаточно *больших числах* n правая часть полученного неравенства произвольно *мала*. Вместе с ней произвольно мала и оцениваемая ею вероятность отклонения частоты $n^{-1}s$ от вероятности a больше чем на ε. В этом выражается *закон больших чисел* для частоты.

2.4. Следствия

Из доказанного неравенства и правила дополнения для вероятности вытекает эквивалентное неравенство. Условимся его тоже называть *теоремой Бернулли*.

Следствие.

$$P\left\{u : \left| n^{-1} \sum_{1 \leqslant j < n} s_j(u) - a \right| < \varepsilon \right\} \geqslant 1 - \varepsilon^{-2} n^{-1} a(1-a).$$

Замечание. Для каждого $\varepsilon > 0$ при всех достаточно *больших числах* n правая часть полученного неравенства произвольно близка к 1. Вместе с ней произвольно близка к 1 и оцениваемая ею вероятность отклонения *частоты* $n^{-1}s$ от *вероятности* a строго меньше чем на ε. Это другая форма *закона больших чисел* для частот.

Лемма. $a(1-a) \leqslant 4^{-1}$.

Доказательство. Это неравенство эквивалентно неравенству

$$(2a-1)^2 \geqslant 0$$

и поэтому верно для каждого вещественного числа a.

Из этой леммы и теоремы Бернулли вытекает еще одно

Следствие. $P\{u : |n^{-1}s(u) - a| \geqslant \varepsilon\} \leqslant (4n)^{-1}\varepsilon^{-2}$.

Замечание. Это неравенство дает оценку для вероятности значительного отклонения частоты $n^{-1}s$ от вероятности a, *не зависящую от вероятности* a. Это часто бывает существенно.

3.4. Пример. Подбрасывание монеты

Как и в § 1, рассмотрим опыт, состоящий из последовательности n подбрасываний симметричной монеты и описываемый *моделью Бернулли n испытаний с вероятностью успеха $a=1/2$.*

Условимся в используемой модели считать события, вероятность которых меньше $a=10^{-3}$, практически невозможными. Используя теорему Бернулли, оценим число n подбрасываний монеты, при котором отклонение частоты герба от вероятности $1/2$ большее $\varepsilon=0{,}0005$, было бы практически невозможно. Из теоремы Бернулли следует, что для этого достаточно, чтобы

$$n \geqslant 4^{-1}\alpha^{-1}\varepsilon^{-2}=4^{-1}10^{3}25^{-1}10^{8}=10^{9}.$$

Таким образом, результат К. Пирсона для $n=24\cdot 10^{3}$ *не противоречит* закону больших чисел.

Замечание. Чувствуется, что миллиард подбрасываний — чересчур большое число. Сказывается грубость оценки в теореме Бернулли.

Задачи, поясняющие использование теоремы Бернулли, собраны в § 2 главы 2 части III.

§ 5*. ЭКСПОНЕНЦИАЛЬНОЕ НЕРАВЕНСТВО

Это неравенство позволяет уточнить теоремы Чебышева и Бернулли.

1.5. Экспоненциальные полиномы

Эти полиномы обладают многими полезными свойствами.

1.1.5. Определение

Для каждого номера $l \geqslant 0$ условимся называть *l-м экспоненциальным полиномом* полином l-й степени, k-й коэффициент которого равен числу $(k!)^{-1}(0 \leqslant k \leqslant l)$. Будем обозначать l-й экспоненциальный полином символом e_l.

Таким образом, по определению,

$$e_l(x)=\sum_{0\leqslant k\leqslant l}\frac{1}{k!}x^k$$

для каждого числа x.

В частности,

$$e_0(x)=1,$$

$$e_1(x)=1+x,$$

$$e_2(x)=1+x+\frac{1}{2}x^2,$$

$$e_3(x)=1+x+\frac{1}{2}x^2+\frac{1}{6}x^3.$$

для каждого числа x.

2.1.5. Формула для произведения

Произведение значений экспоненциального полинома равна его *значению для суммы* аргументов плюс некоторый остаток.

Пример. Для каждых вещественных чисел x_1 и x_2 верны равенства

$$e_2(x_1)e_2(x_2) = (1 + x_1 + (1/2)x_1^2)(1 + x_2 + (1/2)x_2^2) =$$

$$= 1 + (x_2 + x_1) + ((1/2)x_2^2 + x_1x_2 + (1/2)x_1^2) + ((1/2)x_1x_2^2 + (1/2)x_1^2x_2) +$$

$$+ (1/4)x_1^2x_2^2 = e_2(x_1 + x_2) + r_2(x_1, x_2), \quad r_2(x_1, x_2) =$$

$$= ((1/2)x_1x_2^2 + (1/2)x_1^2x_2) + (1/4)x_1^2x_2^2.$$

Значение $e_2(x_1+x_2)$ складывается из произведений, получающихся при перемножении значений для x_1 и x_2 членов полинома e_2, сумма показателей степеней для которых меньше 2:

$$0+0=0,\ 0+1=1+0=1,\ 0+2=1+1=2+0=2.$$

Остаток $r_2(x_1, x_2)$ складывается из произведений, получающихся при перемножении значений для x_1 и x_2 тех членов полиномов, сумма показателей степеней для которых строго больше 2:

$$1+2=2+1=3, \quad 2+2=4.$$

Рассмотрим номер $n \geqslant 1$, семейство чисел $x_1, \ldots, x_n$, сумму $x_1 + \ldots + x_n$ этих чисел и произведение $e_1(x_1), \ldots, e_l(x_n)$ значений $e_l(x_1), \ldots, e_l(x_n)$ экспоненциального полинома e_l ($l \geqslant 0$). Обозначим символом $r_l(x_1, \ldots, x_n)$ сумму всех произведений, получающихся при перемножении значений для $x_1, \ldots, x_n$ тех членов полинома e_l, сумма $k_1 + \ldots + k_n = k$ показателей $k_1, \ldots, k_n$ степеней для которых строго больше l:

$$r_l(x_1, \ldots, x_n) = \sum_{l < k \leqslant nl} \sum_{k_1 + \ldots + k_n = k} \frac{1}{k_1! \ldots k_n!} x_1^{k_1} \ldots x_n^{k_n}.$$

При таких обозначениях формулу для произведения значений экспоненциального полинома выражает

Лемма 1. $e_l(x_1) \ldots e_l(x_n) = e_l(x_1 + \ldots + x_n) + r_l(x_1 + \ldots + x_n)$.

Доказательство. Из дистрибутивности умножения относительно сложения вытекает, что произведение $e_l(x_1) \ldots e_l(x_n)$ сумм $e_l(x_1), \ldots, e_l(x_n)$ равно сумме произведений их слагаемых:

$$e_l(x_1) \ldots e_l(x_n) = \left(\sum_{0 \leqslant k_1 \leqslant l} \frac{1}{k_1!} x_1^{k_1} \right) \ldots \left(\sum_{0 \leqslant k_n \leqslant l} \frac{1}{k_n!} x_n^{k_n} \right) =$$

$$= \sum_{0 \leqslant k_1, \ldots, k_n \leqslant l} \frac{1}{k_1! \ldots k_n!} x_1^{k_1} \ldots x_n^{k_n}.$$

Вследствие ассоциативности сложения эти произведения можно сгруппировать по значениям суммы $k_1+\ldots+k_n=k$ показателей $k_1, \ldots, k_n$ степеней чисел $x_1, \ldots, x_n$:

$$\sum_{0\leqslant k_1,\ldots,k_n\leqslant l} \frac{1}{k_1!\ldots k_n!} x_1^{k_1} \ldots x_n^{k_n} = \sum_{0\leqslant k\leqslant nl} \sum_{\substack{k_1+\ldots+k_n=k \\ (k_1,\ldots,k_n\leqslant l)}} \frac{1}{k_1!\ldots k_n!} x_1^{k_1} \ldots x_n^{k_n}.$$

Если $0\leqslant k\leqslant l$ и $k_1+\ldots+k_n=k$, то $0\leqslant k_1, \ldots, k_n\leqslant l$. Поэтому

$$\sum_{0\leqslant k\leqslant nl} \sum_{\substack{k_1+\ldots+k_n=k \\ (k_1,\ldots,k_n\leqslant l)}} \frac{1}{k_1!\ldots k_n!} x_1^{k_1} \ldots x_n^{k_n} =$$

$$= \sum_{0\leqslant k\leqslant l} \sum_{k_1+\ldots+k_n=k} \frac{1}{k_1!\ldots k_n!} x_1^{k_1} \ldots x_n^{k_n} +$$

$$+ \sum_{l<k\leqslant nl} \sum_{\substack{k_1+\ldots+k_n=k \\ (k_1,\ldots,k_n\leqslant l)}} \frac{1}{k_1!\ldots k_n!} x_1^{k_1} \ldots x_n^{k_n}.$$

С другой стороны, используя доказанную в § 5 главы 2 части I раздела I *полиномиальную формулу*, убеждаемся в том, что:

$$e(x_1+\ldots+x_n) = \sum_{0\leqslant k\leqslant l} \frac{1}{k!}(x_1+\ldots+x_n)^k =$$

$$= \sum_{0\leqslant k\leqslant l} \frac{1}{k!} \sum_{k_1+\ldots+k_n=k} \frac{k!}{k_1!\ldots k_n!} x_1^{k_1} \ldots x_n^{k_n} =$$

$$= \sum_{0\leqslant k\leqslant l} \sum_{k_1+\ldots+k_n=k} \frac{1}{k_1!\ldots k_n!} x_1^{k_1} \ldots x_n^{k_n}.$$

Из полученных равенств следует доказываемое равенство.

3.1.5. Оценка для разности

Сумма членов с достаточно *большими* номерами в экспоненциальном полиноме произвольно *мала*.

Рассмотрим произвольные номер l, номер m и число x **такие,** что $m\geqslant l\geqslant 2|x|$.

Лемма 2.

$$|e_m(x)-e_l(x)| \leqslant 2\frac{|x|^{l+1}}{(l+1)!} \quad (m\geqslant l\geqslant 2|x|).$$

Доказательство. Используя неравенство треугольника, получаем:

$$|e_m(x)-e_l(x)| = \left|\sum_{l<k\leqslant m} \frac{1}{k!}x^k\right| \leqslant \sum_{l<k\leqslant m} \frac{1}{k!}|x|^k.$$

Из условия $l \geqslant 2|x|$ следует, что

$$\frac{(l+1)!}{k!}|x|^{k-l-1} \leqslant \frac{1}{(l+2)\ldots k}|x|^{k-l-1} \leqslant \frac{|x|^{k-l-1}}{(l+2)^{k-l-1}} \leqslant (1/2)^{k-l-1}$$

для каждого номера $k > l+1$. Поэтому

$$\sum_{l<k\leqslant m} \frac{1}{k!}|x|^k = \frac{|x|^{l+1}}{(l+1)!} \sum_{l<k\leqslant m} \frac{(l+1)!}{k!}|x|^{k-l-1} \leqslant$$

$$\leqslant \frac{|x|^{l+1}}{(l+1)!} \sum_{l<k\leqslant m} (1/2)^{k-l-1} \leqslant 2\frac{|x|^{l+1}}{(l+1)!}.$$

Из полученных неравенств следует доказываемое

Замечание. Для достаточно *больших* номеров l правая часть доказанного неравенства произвольно *мала*. Это вытекает из того, что для каждого номера $k \geqslant 2|x|$ верно неравенство $k^{-1}|x| \leqslant 2^{-1}$ и, значит, рассматриваемая правая часть меньше произведения некоторого числа на $2^{-(l-2|x|)}$.

4.1.5. Оценка для остатка

Полученный результат позволяет оценить остаток в формуле для произведения значений экспоненциального полинома.

Лемма 3.

$$|r_l(x_1, \ldots, x_n)| \leqslant 2\frac{(|x_1| + \ldots + |x_n|)^{l+1}}{(l+1)!} \quad (l \geqslant 2(|x_1| + \ldots + |x_n|)).$$

Доказательство. Используя неравенство треугольника, добавляя недостающие слагаемые с показателями степеней $k_1, \ldots, k_n$, некоторые из которых больше l, получаем:

$$|r_l(x_1, \ldots, x_n)| \leqslant \sum_{l<k\leqslant nl} \sum_{\substack{k_1+\ldots+k_n=k \\ (k_1,\ldots,k_n\leqslant l)}} \frac{1}{k_1!\ldots k_n!}|x_1|^{k_1}\ldots|x_n|^{k_n} \leqslant$$

$$\leqslant \sum_{l<k\leqslant nl} \sum_{k_1+\ldots+k_n=k} \frac{1}{k_1!\ldots k_n!}|x_1|^{k_1}\ldots|x_n|^{k_n}.$$

Вследствие полиномиальной формулы отсюда вытекает, что

$$|r_l(x_1, \ldots, x_n)| \leqslant \sum_{l<k\leqslant nl} \frac{1}{k!}(|x_1| + \ldots + |x_n|)^k =$$

$$= e_{nl}(|x_1| + \ldots + |x_n|) - e_l(|x_1| + \ldots + |x_n|).$$

По лемме 2

$$e_{nl}(|x_1| + \ldots + |x_n|) - e_l(|x_1| + \ldots + |x_n|) \leqslant 2\frac{(|x_1| + \ldots + |x_n|)^{l+1}}{(l+1)!},$$

если $l \geqslant 2(|x_1| + \ldots + |x_n|)$.

Лемма 3 доказана.

Пример. Для каждых вещественных чисел x_1 и x_2 верны неравенства

$$|r_2(x_1, x_2)| \leqslant ((1/2)|x_1|\,|x_2|^2+(1/2)|x_1|^2|x_2|)+(1/4)|x_1|^2|x_2|^2 \leqslant$$
$$\leqslant ((1/6)|x_2|^3+(1/2)|x_1|\,|x_2|^2+(1/2)|x_1|^2|x_2|+(1/6)|x_1|^3)+$$
$$+((1/24)|x_2|^4+(1/6)|x_1|\,|x_2|^3+(1/4)|x_1|^2|x_2|^2+(1/6)|x_1|^3|x_2|+$$
$$+(1/24)|x|^4).$$

Следовательно,

$$|r_2(x_1, x_2)| \leqslant \frac{1}{3!}(|x_1|+|x_2|)^3 + \frac{1}{4!}(|x_1|+|x_2|)^4$$

и, значит,

$$|r_2(x_1, x_2)| \leqslant 2\frac{(|x_1|+|x_2|)^3}{3!},$$

если $|x_1|+|x_2| \leqslant 1$.

5.1.5. Положительность значений

Формула для произведения значений и полученная оценка остатка позволяют доказать, что для каждого *вещественного* числа x и достаточно большого номера l значение e_l *строго положительно*. Для положительных чисел x это следует из определений. Рассмотрим произвольное строго отрицательное число x.

Лемма 4. $e_l(x) > 0 \quad (l > (4|x|+1)(\log_2|x|+1))$.

Доказательство. Из определения экспоненциальных полиномов вытекает, что

$$e_l(0)=1,\ e_l(|x|) \geqslant 1 \ (l \geqslant 0).$$

По формуле для произведения значений экспоненциального полинома

$$e_l(x)e_l(-x)=e_l(x-x)+r_l(x, -x)=1+r_l(x, -x).$$

Так как

$$e_l(-x)=e_l(|x|)>0,$$

то отсюда следует, что

$$e_l(x)>0,$$

если

$$1+r_l(x, -x) \geqslant 1-|r_l(x, -x)|>0.$$

Из следствия 1 вытекает, что

$$|r_l(x, -x)| \leqslant 2\frac{(2|x|)^{l+1}}{(l+1)!},$$

если $l \geqslant 4|x|$.

Рассмотрим номер l_0, определяемый неравенствами

$$4|x| \leqslant l_0 < 4|x|+1.$$

Для каждого номера $l > l_0$ имеем:

$$2\frac{(2|x|)^{l+1}}{(l+1)!} = 2\frac{(2|x|)^{l_0}}{l_0!}\frac{(2|x|)^{l+1-l_0}}{(l_0+1)\ldots(l+1)} \leqslant 4\frac{(2|x|)^{l_0}}{2^{l_0}}\frac{(2|x|)^{l+1-l_0}}{(l_0+1)^{l+1-l_0}} =$$

$$= 4|x|^{l_0}\left(\frac{2|x|}{l_0+1}\right)^{l+1-l_0} < 4|x|^{l_0}2^{-(l+1-l_0)}.$$

Из полученных неравенств вытекает, что

$$|r_l(x, -x)| < 1,$$

если

$$4|x|^{l_0}2^{-(l+1-l_0)} \leqslant 1.$$

Это неравенство эквивалентно неравенству

$$2 + l_0\log_2|x| - (l+1-l_0) \leqslant 0,$$

которое, в свою очередь, эквивалентно неравенству

$$l \geqslant l_0(\log_2|x|+1)+1.$$

Из условия следует, что такое неравенство верно. Значит,

$$|r_l(x, -x)| < 1,\ 1 + r_l(x, -x) > 0,\ e_l(x) > 0,$$

что и требовалось доказать.

2.5. Определения и обозначения

В этом пункте рассматриваются *центрированные* и *нормированные* суммы случайных переменных и их композиция с экспоненциальными полиномами.

1.2.5. Центрированные и нормированные суммы

Рассмотрим семейство $(f_j)_{1\leqslant j\leqslant n}$ независимых случайных переменных f_j на множестве U исходов конечного вероятностного пространства $(U, \mathcal{A}; p, P)$. Обозначим средние и стандартные отклонения случайных переменных f_j соответственно символами a_j и b_j:

$$E(f_j) = a_j, \quad D(f_j) = b_j^2 \quad (1 \leqslant j \leqslant n).$$

Положим:

$$a = \sum_{1\leqslant j\leqslant n} a_j, \quad b = \Big(\sum_{1\leqslant j\leqslant n} b_j^2\Big)^{1/2}.$$

Будем предполагать, что хотя бы одна из случайных переменных f_j невырожденная и, следовательно, $b > 0$.

Рассмотрим семейство $(g_j)_{1\leqslant j\leqslant n}$ случайных переменных

$$g_j = b^{-1}(f_j - a_j) \quad (1 \leqslant j \leqslant n)$$

и их сумму

$$g = \sum_{1\leqslant j\leqslant n} g_j.$$

Из независимости случайных переменных f_j следует независимость случайных переменных g_j. В самом деле, для каждых значений y_j переменных g_j равенства $g_j(u)=y_j$ эквивалентны равенствам $f_j(u)=x_j$ для $x_j=by_j+a_j$ и, значит,

$$\{u: g_j(u)=y_j\}=g_j^{-1}(y_j)=f_j^{-1}(x_j)=\{u: f_j(u)=x_j\} \quad (1\leqslant j\leqslant n).$$

Отсюда и из независимости случайных переменных f_j вытекает (лемма 1 пункта 3 § 3 главы 1), что

$$P\left(\bigcap g_j^{-1}(y_j)\right)=P\left(\bigcap f_j^{-1}(x_j)\right)=\Pi P\left(f_j^{-1}(x_j)\right)=\Pi P\left(g_j^{-1}(y_j)\right).$$

Следовательно (та же лемма), случайные переменные g_j независимы $(1\leqslant j\leqslant n)$.

Используя независимость переменных g_j и правила вычисления среднего и дисперсии, нетрудно вычислить среднее и дисперсию рассматриваемой случайной переменной g.

Имеем:

$$E(g_j)=b^{-1}E(f_j-a_j)=b^{-1}(E(f_j)-a_j)=b^{-1}(a_j-a_j)=0,$$

$$D(g_j)=b^{-2}D(f_j-a_j)=b^{-2}D(f_j)=b^{-2}b_j^2$$

для каждого номера $j=1,\ldots,n$.

Следовательно,

$$E(g)=\sum_{1\leqslant j\leqslant n}E(g_j)=0.$$

А так как случайные переменные g_j независимы, то

$$D(g)=\sum_{1\leqslant j\leqslant n}D(g_j)=b^{-2}\sum_{1\leqslant j\leqslant n}b_j^2=b^{-2}b^2=1.$$

Таким образом,

$$E(g)=0,\ D(g)=1.$$

В связи с этим говорят, что сумма g *центрирована* и *нормирована*.

2.2.5. Композиция с полиномом

Для каждых чисел $\varepsilon>0$ и $t>0$ рассмотрим случайные переменные tg_j и $t(g-\varepsilon)$ со значениями соответственно $tg_j(u)$ и $t(g(u)-\varepsilon)$ для каждого исхода u. Композиции этих случайных переменных с l-м экспоненциальным полиномом e_l обозначим символами h_j и h:

$$h_j=e_l\circ(tg_j)=\sum_{0\leqslant k\leqslant l}\frac{1}{k!}t^kg_j^k,$$

$$h=e_l\circ(t(g-\varepsilon))=\sum_{0\leqslant k\leqslant l}\frac{1}{k!}t^k(g-\varepsilon)^k.$$

Из независимости случайных переменных g_j следует независимость случайных переменных $h_j(1\leqslant j\leqslant n)$. В самом деле, для

каждых значений z_j переменных h_j равенства $h_j(u)=z_j$ верны, если и только если $g_j(u)\in C_j$ для некоторых событий C_j, и значит,

$$\{u: h_j(u) = z_j\} = h_j^{-1}(z_j) = g_j^{-1}(C_j) = \{u: g_j(u) \in C_j\}$$

$(1\leqslant j\leqslant n)$. Отсюда и из независимости случайных переменных g_j вытекает (теорема 1 пункта 3 § 3 главы 1), что

$$P\left(\bigcap h_j^{-1}(z_j)\right) = P\left(\bigcap g_j^{-1}(C_j)\right) = \Pi P\left(g_j^{-1}(C_j)\right) = \Pi P\left(h_j^{-1}(z_j)\right).$$

Следовательно (лемма 1 того же пункта), случайные переменные *h_j независимы* $(1\leqslant j\leqslant n)$.

3.5. Экспоненциальное неравенство

Это неравенство оценивает вероятность *значительного* отклонения центрированной и нормированной суммы независимых случайных переменных от нуля. Оценка производится с помощью *экспоненциального полинома.*

1.3.5. Неравенство для вероятности

Используя положительность экспоненциального полинома, можно оценить вероятность значительного отклонения суммы g от нуля в положительную сторону с помощью среднего для h.

Рассмотрим число $c>0$, ограничивающее абсолютные значения всех случайных переменных g_j:

$$|g_j(u)|\leqslant c.$$

для всех исходов u и номеров $j=1, \ldots, n$. Заметим, что

$$|g(u)| \leqslant \sum_{1\leqslant j\leqslant n} |g_j(u)| \leqslant nc$$

для всех исходов u.

Будем предполагать, что:

(1) $t\leqslant 1/c$,

(2) $l\geqslant 1/c^2+(4(n+\varepsilon/c)+1)(\log_2(n+\varepsilon/c)+1)$.

При этих предположениях из полученных результатов вытекает

Предложение 1. $P\{u: g(u)\geqslant\varepsilon\}\leqslant E(h)$.

Доказательство. Рассмотрим событие

$$A=\{u: g(u)\geqslant\varepsilon\}=\{u: g(u)-\varepsilon\geqslant 0\},$$

описывающее значительное отклонение суммы g от нуля в положительную сторону.

Из определений следует, что

$$h(u)=e_l(t(g(u)-\varepsilon))\geqslant 1 \quad (u\in A).$$

Из условий 1 и 2 следует, что

$$l > (4t(nc+\varepsilon)+1)(\log_2(t(nc+\varepsilon))+1) \geqslant$$
$$\geqslant (4t|g(u)-\varepsilon|+1)(\log_2(t|g(u)-\varepsilon|)+1)$$

для каждого исхода u.

По лемме 4 отсюда вытекает, что

$$h(u)=e_l(t(g(u)-\varepsilon))>0 \quad (u\in U).$$

Значит,

$$i_A \leqslant h.$$

По правилу индикатора и правилу неравенства для средних отсюда следует, что

$$P(A)=E(i_A)\leqslant E(h).$$

Это и требовалось доказать.

2.3.5. Предварительные неравенства для средних

Чтобы получить неравенство для среднего случайной переменной h, предварительно оценим средние для случайных переменных h_j. По-прежнему будем предполагать, что верны неравенства 1 и 2.

Лемма 5.

$$E(h_j)\leqslant e_l\left(\frac{b_j^2}{b}\, t^2/2\left(1+\frac{tc}{2}\right)\right) \quad (1\leqslant j\leqslant n).$$

Доказательство. Рассмотрим произвольный номер $j\,(1\leqslant j\leqslant n)$. Вследствие правила линейности для среднего

$$E(h_j)=\sum_{0\leqslant k\leqslant l}\frac{t^k}{k!}E\left(g_j^k\right).$$

Так как

$$E\left(g_j^0\right)=1, \quad E\left(g_j^1\right)=0,$$

то

$$E(h_j)=1+\sum_{2\leqslant k\leqslant l}\frac{t^k}{k!}E\left(g_j^k\right).$$

Заметим, что

$$g_j^k\leqslant c^{k-2}g_j^2 \quad (k\geqslant 2).$$

Так как

$$E\left(g_j^2\right)=D\left(g_j\right)=\frac{b_j^2}{b^2},$$

то

$$E\left(g_j^k\right)\leqslant c^{k-2}E\left(g_j^2\right)=c^{k-2}\frac{b_j^2}{b^2} \quad (k\geqslant 2).$$

Кроме того,

$$c^{k-2}\,\frac{t^k}{k!}\leqslant c^{k-2}\,\frac{t^k}{2\cdot 3^{k-2}}=\frac{t^2}{2}\left(\frac{tc}{3}\right)^{k-2}\quad (k\geqslant 2).$$

Используя эти неравенства, получаем:

$$\sum_{2\leqslant k\leqslant l}\frac{t^k}{k!}\,E\left(g_j^k\right)\leqslant\frac{b_j^2}{b^2}\,\frac{t^2}{2}\sum_{2\leqslant k\leqslant l}\left(\frac{tc}{3}\right)^{k-2}.$$

Отсюда и из условия 1 следует, что

$$\sum_{2\leqslant k\leqslant l}\frac{t^k}{k!}\,E\left(g_j^k\right)\leqslant\frac{b_j^2}{b_2}\,\frac{t^2}{2}\,\frac{1}{1-3^{-1}tc}.$$

Если $tc\leqslant 1$, то

$$\frac{1}{1-3^{-1}tc}=1+\frac{3^{-1}tc}{1-3^{-1}tc}\leqslant 1+\frac{3^{-1}tc}{1-3^{-1}}=1+\frac{tc}{2}.$$

Поэтому

$$\sum_{2\leqslant k\leqslant l}\frac{t^k}{k!}\,E\left(g_j^k\right)\leqslant\frac{b_j^2}{b^2}\,\frac{t^2}{2}\left(1+\frac{tc}{2}\right).$$

Следовательно,

$$E\left(h_j\right)\leqslant 1+\frac{b_j^2}{b^2}\,\frac{t^2}{2}\left(1+\frac{tc}{2}\right).$$

Из определения экспоненциального полинома e_l вытекает, что для каждого номера $l>0$ и каждого числа $x\geqslant 0$ верно неравенство

$$1+x\leqslant e_l(x).$$

Из условия 2 следует, что $l>0$. Поэтому из полученного неравенства для $E(h_j)$ вытекает доказываемое.

3.3.5. Неравенство для среднего

Используя формулу для произведения значений экспоненциальных полиномов, правило произведения для средних и имеющиеся оценки, можно получить неравенство для среднего случайной переменной h.

Предполагается, что верны неравенства 1 и 2.

Предложение 2.

$$E(h)\leqslant e_l\left(-t\varepsilon+\frac{t^2}{2}\left(1+\frac{tc}{2}\right)\right)+r_l,$$

$$r_l\leqslant\frac{2}{(l+1)!}\left(\left(n+\frac{\varepsilon}{c}\right)^{l+1}+\left(\frac{1}{c^2}+\frac{\varepsilon}{c}\right)^{l+1}\right).$$

Доказательство. Из определения случайной переменной h и формулы для произведения значений экспоненциального полинома вытекает, что

$$h(u)=e_l(t(g(u)-\varepsilon))=e_l(tg_1(u)+\ldots+tg_n(u)-$$
$$-t\varepsilon)=e_l(-t\varepsilon)e_l(tg_1(u))\ldots e_l(tg_n(u))-$$
$$-r_l(tg_1(u),\ldots,tg_n(u),-t\varepsilon)=e_l(-t\varepsilon)h_1(u)\ldots h_n(u)-$$
$$-r_l(tg_1(u),\ldots,tg_n(u),-t\varepsilon)$$

для каждого исхода u. По условию 1

$$|tg_1(u)|+\ldots+|tg_n(u)|+|-t\varepsilon|\leqslant ntc+t\varepsilon\leqslant n+\varepsilon/c.$$

По условию 2 и лемме 3 отсюда следует, что

$$|r_l(tg_1(u),\ldots,tg_n(u),-t\varepsilon|\leqslant\alpha=\frac{2}{(l+1)!}\left(n+\frac{\varepsilon}{c}\right)^{l+1}.$$

Значит,

$$h(u)\leqslant e_l(-t\varepsilon)h_1(u)\ldots h(u)+\alpha$$

для каждого исхода u.

Так как случайные переменные $h_1,\ldots,h_n$ независимы, то, используя правила неравенства, линейности и умножения для среднего, выводим из полученного неравенства, что

$$E(h)\leqslant e_l(-t\varepsilon)E(h_1)\ldots E(h_n)+\alpha.$$

По лемме 5

$$e_l(-t\varepsilon)E(h_1)\ldots E(h_n)\leqslant e_l(-t\varepsilon)e_l(c_1)\ldots e_l(c_n),$$
$$c_1=\frac{b_1^2}{b^2}\frac{t^2}{2}\left(1+\frac{tc}{2}\right),\ldots,c_n=\frac{b_n^2}{b^2}\frac{t^2}{2}\left(1+\frac{tc}{2}\right).$$

Используя формулу для произведения значений экспоненциального полинома и равенство $b_1^2+\ldots+b_n^2=b^2$, убеждаемся в том, что правая часть неравенства равна

$$e_l\left(-t\varepsilon+\frac{t^2}{2}\left(1+\frac{tc}{2}\right)\right)+r_l(c_1,\ldots,c_n,-t\varepsilon).$$

По условию 1

$$|c_1|+\ldots+|c_n|+|-t\varepsilon|=\frac{t^2}{2}\left(1+\frac{tc}{2}\right)+t\varepsilon\leqslant\frac{1}{c^2}+\frac{\varepsilon}{c}.$$

По условию 2 и лемме 3 отсюда следует, что

$$|r_l(c_1,\ldots,c_n,-t\varepsilon)|\leqslant\beta=\frac{2}{(l+1)!}\left(\frac{1}{c^2}+\frac{\varepsilon}{c}\right)^{l+1}.$$

Значит,

$$E(h)\leqslant e_l\left(-t\varepsilon+\frac{t^2}{2}\left(1+\frac{tc}{2}\right)\right)+(\alpha+\beta),$$

что и требовалось доказать.

Замечание. Для достаточно больших номеров l остаток r_l произвольно мал.

4.3.5. Экспоненциальное неравенство

Из полученных для вероятности и среднего неравенств следует основное неравенство, оценивающее вероятность *значительного* отклонения от нуля центрированной и нормированной суммы g независимых случайных переменных $g_j (1 \leqslant j \leqslant n)$, абсолютные значения которых ограничены сверху числом $c > 0$.

Для каждых числа $\varepsilon > 0$, достаточно малого числа $t > 0$ и достаточно большого номера l, для которых

$$t \leqslant 1/c, \quad l \geqslant \frac{1}{c^2} + \left(4\left(n + \frac{\varepsilon}{c}\right) + 1\right)\left(\log_2\left(n + \frac{\varepsilon}{c}\right) + 1\right),$$

верно

Предложение 3.

$$P\{u : g(u) \geqslant \varepsilon\} \leqslant e_l\left(-t\varepsilon + \frac{t^2}{2}\left(1 + \frac{tc}{2}\right)\right) + r_l,$$

$$r_l \leqslant \frac{2}{(l+1)!}\left(\left(n + \frac{\varepsilon}{c}\right)^{l+1} + \left(\frac{1}{c^2} + \frac{\varepsilon}{c}\right)^{l+1}\right).$$

Доказательство. Это предложение следует из предложений 1 и 2.

Если $\varepsilon \leqslant c^{-1}$, то условию $t \leqslant c^{-1}$ удовлетворяет число $t = \varepsilon$. В этом случае

$$-t\varepsilon + \frac{t^2}{2}\left(1 + \frac{tc}{2}\right) = -\varepsilon^2 + \frac{\varepsilon^2}{2}\left(1 + \frac{\varepsilon c}{2}\right) = -\frac{\varepsilon^2}{4}(2 - \varepsilon c) \leqslant -\frac{\varepsilon^2}{4}.$$

Если $\varepsilon > c^{-1}$, то ближайшим к ε числом, удовлетворяющим условию $t \leqslant c^{-1}$, является число $t = c^{-1}$. В этом случае

$$-t\varepsilon + \frac{t^2}{2}\left(1 + \frac{tc}{2}\right) = -\frac{\varepsilon}{c} + \frac{1}{2c^2}\left(1 + \frac{1}{2}\right) = -\frac{\varepsilon}{c} + \frac{3}{4}\frac{1}{c^2} \leqslant$$

$$\leqslant -\frac{\varepsilon}{c} + \frac{3}{4} \cdot \frac{\varepsilon}{c} = -\frac{1}{4} \cdot \frac{\varepsilon}{c} = -\frac{\varepsilon^2}{4} \cdot \frac{1}{\varepsilon c}.$$

Обозначим наименьшее из чисел 1 и $(\varepsilon c)^{-1}$ символом $\varphi(\varepsilon c)$:

$$\varphi(\varepsilon c) = \min\{1, (\varepsilon c)^{-1}\}.$$

Если $\varepsilon c \leqslant 1$, то $\varphi(\varepsilon c) = 1$, а если $\varepsilon c > 1$, то $\varphi(\varepsilon c) = (\varepsilon c)^{-1} < 1$. Из полученных для рассматривавшейся суммы неравенств следует, что

$$-\varepsilon t + \frac{t^2}{2}\left(1 + \frac{tc}{2}\right) \leqslant -\frac{\varepsilon^2}{4} \cdot \varphi(\varepsilon c),$$

если $t = \varepsilon \leqslant c^{-1}$ или $\varepsilon > c^{-1} = t$.

При сделанных предположениях и введенных обозначениях из предположения 3 вытекает.

Следствие 1.

$$P\{u:g(u)\geqslant\varepsilon\}\leqslant e_l(-\varepsilon^2/4\cdot\varphi(\varepsilon c))+r_l.$$

Замечание. Проведенные рассуждения показывают, что это неравенство верно также, если

$$\varphi(\varepsilon c)=2-\varepsilon c\ \ (\varepsilon c\leqslant 1),\ \varphi(\varepsilon c)=(\varepsilon c)^{-1}(\varepsilon c>1).$$

Полученное неравенство оценивает вероятность значительного отклонения переменной g от нуля в положительную сторону. При тех же предположениях из него легко вывести аналогичные оценки для вероятностей значительного отклонения от нуля в отрицательную сторону и значительного отклонения от нуля в любую сторону.

Следствие 2.

$$P\{u:g(u)\leqslant-\varepsilon\}\leqslant e_l(-\varepsilon^2/4\cdot\varphi(\varepsilon c))+r_l.$$

Доказательство. Это неравенство вытекает из неравенства следствия 1 для суммы $-g$ случайных переменных $-g_j$, полученных из случайных переменных $-f_j$.

Из независимости случайных переменных f_j следует независимость случайных переменных $-f_j$ и, кроме того,

$$E(-f_j)=-E(f_j)=-a_j,$$
$$D(-f_j)=D(f_j)=b_j^2\qquad\qquad(1\leqslant j\leqslant n).$$

Поэтому

$$-g_j=-b^{-1}(f_j-a_j)=b^{-1}((-f_j)-(-a_j))\quad(1\leqslant j\leqslant n).$$

Следовательно, для случайной переменной $-g$ верно неравенство

$$P\{u:-g(u)\geqslant\varepsilon\}\leqslant e_l(-\varepsilon^2/4\cdot\varphi(\varepsilon c))+r_l.$$

Так как

$$\{u:-g(u)\geqslant\varepsilon\}=\{u:g(u)\leqslant-\varepsilon\},$$

то отсюда вытекает доказываемое неравенство.

Следствие 3.

$$P\{u:|g(u)|\geqslant\varepsilon\}\leqslant 2[e_l(-\varepsilon^2/4\cdot\varphi(\varepsilon c))+r_l].$$

Доказательство. Так как

$$\{u:|g(u)|\geqslant\varepsilon\}=\{u:g(u)\geqslant\varepsilon\}+\{u:g(u)\leqslant-\varepsilon\}$$

и, значит,

$$P\{u:|g(u)|\geqslant\varepsilon\}=P\{u:g(u)\geqslant\varepsilon\}+P\{u:g(u)\leqslant-\varepsilon\},$$

то это неравенство вытекает из неравенств следствий 1 и 2.

Условимся *экспоненциальным неравенством* называть неравенство следствия 3. Его можно считать основным результатом данного параграфа.

4.5. Уточнение теорем Чебышева и Бернулли

Экспоненциальное неравенство позволяет уточнить теоремы Чебышева и Бернулли.

1.4.5. Уточнение теоремы Чебышева

Как и в § 3, рассмотрим семейство $(f_j)_{1\leqslant j\leqslant n}$ *независимых* и *одинаково распределенных невырожденных* случайных переменных f_j на множестве исходов U конечной вероятностной модели $(U, \mathscr{A}; p, P)$. Положим:

$$E(f_j)=a,\ D(f_j)=b^2\quad (b>0,\ 1\leqslant j\leqslant n).$$

В соответствии с предыдущим пунктом рассмотрим также семейство $(g_j)_{1\leqslant j\leqslant n}$ случайных переменных g_j со значениями

$$g_j(u)=n^{-1/2}b^{-1}(f_j(u)-a)\quad (u\in U,\ 1\leqslant j\leqslant n)$$

и их сумму g со значениями

$$g(u)=\sum_{1\leqslant j\leqslant n} g_j(u)=n^{1/2}b^{-1}\left(n^{-1}\sum_{1\leqslant j\leqslant n} f_j(u)-a\right)(u\in U).$$

Сумма g *пропорциональна* отклонению арифметического среднего случайных переменных f_j от их среднего a. Следовательно, для оценки вероятности *значительного* такого отклонения можно использовать экспоненциальное неравенство.

Рассмотрим произвольное число $\varepsilon>0$ и число $c>0$, ограничивающее сверху абсолютные значения всех случайных переменных g_j:

$$|g_j(u)|=n^{-1/2}b^{-1}|f_j(u)-a|\leqslant c$$

для каждых исхода u и номера $j=1, \ldots, n$. Рассмотрим также достаточно большой номер

$$l\geqslant\frac{1}{c^2}+\left(4\left(n+\frac{\varepsilon}{c}\right)+1\right)\left(\log_2\left(n+\frac{\varepsilon}{c}\right)+1\right).$$

Для таких чисел ε, c и номера l при сделанных предположениях и введенных обозначениях верна

Теорема Чебышева.

$$P\left\{u:\left|n^{-1}\sum_{1\leqslant j\leqslant n} f_j(u)-a\right|\geqslant n^{-1/2}b\varepsilon\right\}\leqslant$$
$$\leqslant 2\,[e_l(-\varepsilon^2/4\cdot\varphi(\varepsilon c))+r_l].$$

Доказательство. Из определения суммы g вытекает, что

$$\{u:|g(u)|\geqslant\varepsilon\}=\left\{u:n^{1/2}b^{-1}\left|n^{-1}\sum_{1\leqslant j\leqslant n} f_j(u)-a\right|\geqslant\varepsilon\right\}=$$
$$=\left\{u:\left|n^{-1}\sum_{1\leqslant j\leqslant n} f_j(u)-a\right|\geqslant n^{-1/2}b\varepsilon\right\}.$$

Поэтому доказываемое неравенство следует из экспоненциального неравенства, в котором

$$\varphi(\varepsilon c) = \min\{1, (\varepsilon c)^{-1}\},$$

$$r_l \leqslant \frac{2}{(l+1)!}\left(\left(n + \frac{\varepsilon}{c}\right)^{l+1}\left(\frac{1}{c^2} + \frac{\varepsilon}{c}\right)^{l+1}\right).$$

Теорема доказана.

Замечание. Из доказанной раньше теоремы Чебышева следует, что

$$P\left\{u : \left|n^{-1}\sum_{1\leqslant j\leqslant n} f_j(u) - a\right| \geqslant n^{-1/2}b\varepsilon\right\} \leqslant \varepsilon^{-2}.$$

Если число ε достаточно мало, то правые части обоих неравенств больше 1 и эти неравенства тривиальны. Если же число ε и номер l достаточно велики, то правая часть нового неравенства меньше ε^{-2}, и поэтому оно точнее.

Подчеркнем в то же время, что при доказательстве теоремы Чебышева в § 3 предполагалась только *попарная* независимость рассматривавшихся случайных переменных.

Для вычислений, связанных с экспоненциальными полиномами больших степеней, можно использовать имеющиеся таблицы.

В частности, если $l \geqslant 30$, можно пользоваться таблицей

x	1,5	2	2,5	3	3,5	4	5
$2(e_l(-x^2)+r_l)$	$2108 \cdot 10^{-4}$	$3693 \cdot 10^{-4}$	$3860 \cdot 10^{-6}$	$2510 \cdot 10^{-7}$	$9600 \cdot 10^{-9}$	$2300 \cdot 10^{-10}$	$3000 \cdot 10^{-13}$

Эту таблицу интересно сравнить с аналогичной таблицей для вычислений, связанных с неравенством Чебышева:

x	1,5	2	2,5	3	3,5	4	5
x^{-2}	$4444 \cdot 10^{-4}$	$2500 \cdot 10^{-4}$	$1600 \cdot 10^{-4}$	$1111 \cdot 10^{-4}$	$8163 \cdot 10^{-5}$	$6250 \cdot 10^{-5}$	$4000 \cdot 10^{-5}$

Такое сравнение очень убедительно подтверждает бо́льшую точность экспоненциального неравенства.

Пример. Уточненная теорема Чебышева позволяет получить сравнительно более хорошие результаты для примера о точности и числе измерений, рассматривавшегося в пункте 3.3.

Как и в этом пункте, предположим, что $b=1$, $n=10^3$ и $\alpha=10^{-3}$. Дополнительно условимся считать, что возможные отклонения прибора не превосходят 5 единиц:

$$|g_j(u)| = 10^{-3/2}\cdot 1\cdot 5 \leqslant 6^{-1} = c.$$

Используя таблицу, убеждаемся в том, что при $l \geqslant 30$ неравенство

$$2(e_l(-x^2) + r_l) \leqslant \alpha \quad (x = 2^{-1}\varepsilon\varphi^{1/2}(\varepsilon c))$$

верно, если $x=3$. Следовательно,

$$\varepsilon=6 \quad (\varepsilon c=6\cdot 6^{-1}=1,\ \varphi(\varepsilon c)=1).$$

По новой теореме Чебышева отсюда вытекает, что практически достоверная точность измерений

$$\bar{\varepsilon}=n^{-1/2}b\varepsilon=10^{-3/2}\cdot 1\cdot 6<0{,}2$$

(в пункте 3.3 была получена точность 1).

Точно так же, если $b=10^{-1}$, $\bar{\varepsilon}=10^{-1}$ и $\alpha=10^{-2}$, из таблицы видно, что при $l\geqslant 30$ неравенство

$$2(e_l(-x^2)+r_l)\leqslant\alpha \quad (x=2^{-1}\varepsilon\varphi^{1/2}(\varepsilon c))$$

верно, если $x=2{,}5$. Следовательно,

$$\varepsilon=5 \quad (\varepsilon c=5\cdot 6^{-1}<1,\ \varphi(\varepsilon c)=1).$$

По новой теореме Чебышева отсюда вытекает, что число n измерений, практически достоверно обеспечивающее нужную точность $\bar{\varepsilon}$, определяется неравенством

$$n^{-1/2}b\varepsilon\leqslant\bar{\varepsilon}.$$

Значит,

$$n\geqslant\bar{\varepsilon}^{-2}b^2\varepsilon^2=10^2\cdot 10^{-2}\cdot 25=25$$

(в пункте 3.3 было получено неравенство $n\geqslant 100$).

Упражнение. Доказать, что для каждых чисел $\alpha>0$ и $\gamma>0$ существуют число $x_0>0$ и номер l_0 такие, что

$$\alpha e_l(\gamma x)>x^2$$

для каждых числа $x\geqslant x_0$ и номера $l\geqslant l_0$.

Используя это неравенство и вытекающее из формулы для произведения значений экспоненциального полинома равенство

$$e_l(x)e_l(-x)=1+r_l(x,\ -x),$$

можно доказать, что для каждых чисел $\alpha>0$ и $\gamma>0$ существуют число $x_0>0$ и номер l_0 такие, что

$$\alpha e_l(-\gamma x)<x^{-2}$$

для каждых числа $x>x_0$ и номера $l\geqslant l_0$.

2.4.5. Уточнение теоремы Бернулли

Как и в § 4, рассмотрим модель Бернулли n испытаний с вероятностью успеха a и арифметическое среднее $n^{-1}s$ независимых и одинаково распределенных случайных переменных $s_j (1\leqslant j\leqslant n)$. Верны равенства

$$E(s_j)=a,\quad D(s_j)=a(1-a)\leqslant 4^{-1} \quad (1\leqslant j\leqslant n).$$

Отсюда и из новой теоремы Чебышева вытекает новая

Теорема Бернулли.

$$P\left\{u:\left|n^{-1}\sum_{1\leqslant j\leqslant n}s_j(u)-a\right|\geqslant n^{-1/2}\cdot 2^{-1}\cdot\varepsilon\right\}\leqslant$$
$$\leqslant 2\,[e_l(-\varepsilon^2/4\cdot\varphi(\varepsilon c))+r_l].$$

Пример. Уточненная теорема Бернулли позволяет получить сравнительно более хороший результат для примера о подбрасывании монеты, рассматривавшегося в пункте 3.4.

Как и в этом пункте, предположим, что $\alpha=10^{-3}$ и $\bar{\varepsilon}=5\cdot 10^{-4}$. Возможные отклонения можно оценить числом $c=10^{-1}$, считая, что число подбрасываний $n\geqslant 10^2$:

$$|g_j(u)|=n^{-1/2}b^{-1}|s_j(u)-a|\leqslant 10^{-1}\cdot 2\cdot 2^{-1}=10^{-1}.$$

Как и в примере с измерениями в предыдущем пункте, из условия $\alpha=10^{-3}$ получаем $\varepsilon=6$. По новой теореме Бернулли отсюда вытекает, что число n подбрасываний монеты, при котором отклонение частоты герба от вероятности 1/2, большее $\bar{\varepsilon}$, было бы практически невозможным, определяется неравенством

$$n^{-1/2}\cdot 2^{-1}\varepsilon\leqslant\bar{\varepsilon}.$$

Значит,

$$n\geqslant\bar{\varepsilon}^{-2}\cdot 2^{-2}\cdot\varepsilon^2=5^{-2}\cdot 10^8\cdot 4^{-1}\cdot 6^2=36\cdot 10^6$$

(в пункте 3.4 было получено неравенство $n\geqslant 10^9$).

Часть III

ЗАДАЧИ

В этой части рассматриваются несколько десятков задач, связанных с конечными вероятностными моделями и случайными переменными для таких моделей. Определения и факты, необходимые для решения задач, повторяются. Поэтому третью часть книги можно читать независимо от первых двух.

Часть рассматриваемых задач дает представление о самых простых приложениях теории вероятностей. Некоторые из этих задач имеют определенный практический интерес.

Сравнительно более трудные задачи выделены в самостоятельную главу.

Глава 1

КОНЕЧНЫЕ ВЕРОЯТНОСТНЫЕ МОДЕЛИ

В этой главе рассматриваются некоторые задачи, связанные с *конечными вероятностными моделями.*

§ 1. СВОЙСТВА ВЕРОЯТНОСТИ

Конечная вероятностная модель определяется конечным множеством U, элементы u которого называются *исходами*, и положительной нормированной функцией p на множестве U, которая называется *элементарной вероятностью*:

$$p(u) \geqslant 0, \ \Sigma p(u) = 1.$$

Каждая часть A множества исходов U называется *событием*. Сумма $P(A)$ элементарных вероятностей $p(u)$ исходов u, составляющих событие A,— *вероятностью события* A:

$$P(A) = \sum_{u \in A} p(u).$$

Конечная вероятностная модель описывает *опыт* с конечным множеством возможных *результатов*, реализация которых зависит от случая. Исходы описывают возможные результаты опыта. Со-

бития описывают *явления*, определяющиеся этими результатами. Элементарная вероятность исхода является *мерой реализуемости результата*, описываемого этим исходом. Вероятность события служит *мерой реализуемости* явления, описываемого этим событием.

Замечание. Для того, чтобы избежать громоздких оборотов, вместо *результат*, описываемый исходом, говорят *исход*; вместо *явление*, описываемое событием,— *событие*; вместо *мера реализуемости — вероятность.*

Таким образом, термины исход, событие и вероятность употребляются в двух смыслах: содержательном и формальном.

При вычислении вероятностей событий удобно использовать следующие правила:

Правило сложения: $P(A+B)=P(A)+P(B) \quad (AB)=0.$

Правило вычитания: $P(B-A)=P(B)-P(A) \quad (A\subseteq B).$

Правило дополнения: $P(A')=1-P(A).$

Правило объединения: $P(A\cup B)=P(A)+P(B)-P(AB).$

Правило неравенства: $P(A)\leqslant P(B) \quad (A\subseteq B).$

Среди конечных вероятностных моделей выделяются модель Лапласа и модель Бернулли.

1.1. Модель Лапласа

Классическая *модель Лапласа* n равновероятных исходов определяется множеством U из n элементов и элементарной вероятностью p со значением

$$p(u)=1/n$$

для каждого исхода u. В этой модели вероятность $P(A)$ каждого события A равна отношению числа исходов $n(A)$, составляющих событие A, к числу $n(U)$ всех исходов:

$$P(A)=n(A)/n(U).$$

Модель Лапласа описывает опыты с *равновозможными* результатами.

Задача 1. *Игральная кость подбрасывается два раза. Какова вероятность того, что хотя бы один раз появляется шестерка?*

Решение. При решении задачи 10 из § 1 главы 1 было показано, что результаты двукратного подбрасывания кости можно описать множеством U строк $u=u_1u_2$ длины 2, составленных из чисел 1, 2, 3, 4, 5, 6. Число таких строк равно $6^2=36$. Симметричность кости позволяет использовать модель Лапласа $n=36$ равновероятных исходов. Задача сводится к вычислению вероятности $P(C)$ события C, составленного из строк $u=u_1u_2$, для которых $u_1=6$ или $u_2=6$:

$$C=\{61, 62, 63, 64, 65, 66, 16, 26, 36, 46, 56\}.$$

1. Имеем:

$$P(C)=n(C)/n(U)=11/36;$$

2. Дополнение $A=C'$ события C состоит из строк $u=u_1u_2$, для которых $u_1\neq 6$ и $u_1\neq 6$. Число таких строк равно $5^2=25$. Поэтому

$$P(C')=P(A)=5^2/6^2=(5/6)^2.$$

По правилу дополнения,

$$P(C)=P(A')=1-P(A)=1-P(C')=1-(5/6)^2=11/36;$$

3. Событие C равно объединению событий

$$A=\{61, 62, 63, 64, 65, 66\} \text{ и } B=\{16, 26, 36, 46, 56, 66\},$$

описывающих появление шестерки соответственно при первом и втором подбрасываниях.

Имеем:

$$P(A)=6/36,\ P(B)=6/36,\ P(AB)=P(\{66\})=1/36.$$

По правилу объединения,

$$P(C)=P(A\cup B)=P(A)+P(B)-P(AB)=6/36+$$
$$+6/36-1/36=11/36.$$

Задача 2. *В урне находятся 5 шаров, отличающихся только номерами 1, 2, 3, 4, 5. Вынимается наугад выбранный шар и отмечается его номер. Вынутый шар возвращается в урну. После тщательного перемешивания из нее вновь вынимается наугад выбранный шар. Какова вероятность того, что вынимается не один и тот же шар?*

Решение. При решении задачи 11 из § 1 главы 3 части I было показано, что результаты рассматриваемого опыта можно описать множеством U строк $u=u_1u_2$ длины 2, составленных из номеров 1, 2, 3, 4, 5. Число таких строк равно $5^2=25$. Условие о выборе наугад позволяет использовать модель Лапласа $n=25$ равновероятных исходов. Задача сводится к вычислению вероятности $P(A)$ события A, составленного из всех строк $u=u_1u_2$ с *различными* номерами $u_1\neq u_2$. Дополнение A' события A состоит из строк $u=u_1u_2$ с одинаковыми номерами $(u_1=u_2)$:

$$A'=\{11, 22, 33, 44, 55\}.$$

По правилу дополнения,

$$P(A)=1-P(A')=1-n(A')/n(U)=1-5/25=4/5.$$

Задача 3. *В урне находятся 5 шаров, отличающихся только номерами 1, 2, 3, 4, 5. Вынимается наугад выбранный шар. Вынутый шар не возвращается в урну. Вновь вынимается наугад выбранный шар. Какова вероятность того, что номера вынимаемых шаров нечетные или в сумме меньше пяти?*

Решение. При решении задачи 12 из § 1 главы 3 части I было показано, что результаты рассматриваемого опыта можно описать множеством U строк u_1u_2 длины 2, составленных из неравных но-

меров u_1 и u_2 из множества $\{1, 2, 3, 4, 5\}$. Число таких строк равно 20. Условие о выборе наугад позволяет использовать модель Лапласа $n=20$ равновероятных исходов. Задача сводится к вычислению вероятности $P(A \cup B)$ объединения $A \cup B$ событий A и B, составленных из строк $u=u_1u_2 (u_1 \neq u_2)$, для которых соответственно u_1 и u_2 нечетные и $u_1+u_2 \leqslant 5$:

$$A=\{13, 15, 31, 35, 51, 53\}, \ B=\{12, 13, 14, 21, 23, 31, 32, 41\},$$
$$AB=\{13, 31\}.$$

Имеем:

$$P(A)=6/20, \ P(B)=8/20, \ P(AB)=2/20.$$

По правилу объединения,

$$P(A \cup B)=P(A)+P(B)-P(AB)=6/20+8/20-2/20=3/5.$$

Задача 4. *Подбрасываются красная и белая игральные кости. Какова наиболее вероятная сумма числа очков?*

Решение. По аналогии с задачей 1 результаты рассматриваемого опыта можно описать множеством U строк $u=u_1u_2$ длины 2, составленных из чисел 1, 2, 3, 4, 5, 6, и использовать модель Лапласа $n=36$ равновероятных исходов. Задача сводится к вычислению вероятностей $P(A_s)$ событий A_s, составленных из строк $u=u_1u_2$, для которых $u_1+u_2=s$ $(s=2, 3, \ldots, 12)$:

$$A_2=\{11\}, \ A_3=\{12, 21\}, \ A_4=\{13, 22, 31\},$$
$$A_5=\{14, 23, 32, 41\}, \ A_6=\{15, 24, 33, 42, 51\},$$
$$A_7=\{16, 25, 34, 43, 52, 61\}, \ A_8=\{26, 35, 44, 53, 62\},$$
$$A_9=\{36, 45, 54, 63\}, \ A_{10}=\{46, 55, 64\},$$
$$A_{11}=\{56, 65\}, \ A_{12}=\{66\}.$$

Искомые вероятности содержит следующая таблица:

s	2	3	4	5	6	7	8	9	10	11	12
$P(A_s)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Наиболее вероятно, что сумма числа очков окажется равной 7.

Задача 5. *Подбрасываются красная и белая игральные кости. Какова вероятность того, что сумма очков равна: 1) 7 или 11; 2) 2 или 3 или 12; 3) любому другому возможному числу?*

Решение. Используем модель задачи 4. Дело сводится к вычислению вероятностей событий

$$B=A_7+A_{11}, \ П=A_2+A_3+A_{12}, \ H=A_4+A_5+A_6+A_8+A_9+A_{10}.$$

Употребляя *правило сложения* и таблицу задачи 4, получаем:

$$P(A)=P(У_1У_2)+P(H_1H_2)=a^2+(1-a)^2=(0{,}52)^2+(0{,}48)^2 \approx 0{,}51.$$
$$P(H)=3/36+4/36+5/36+5/36+4/36+3/36=6/9.$$

Замечание. Задача 5 имеет отношение к игре *крэпс.* Событие B описывает *выигрыш* в этой игре, Π — *проигрыш,* H — *ничью.* В случае ничьей кости подбрасываются снова и игра продолжается до тех пор, пока не повторится сумма очков, полученная при первом подбрасывании, либо не получится сумма 7. В первом случае игрок выигрывает, во втором — проигрывает.

2.1. Модель Бернулли

Модель Бернулли n *испытаний с вероятностью успеха* a определяется множеством U строк $u=u_1\ldots u_n$ длины n, составленных из *единиц* и *нулей*, и элементарной вероятностью p со значением

$$p(u)=a^{s(u)}(1-a)^{n-s(u)}\ (s(u)=u_1+\ldots+u_n)$$

для каждого исхода u.

Модель Бернулли описывает *опыты*, составленные из *последовательности нескольких одинаковых и независимых испытаний,* каждое из которых имеет *два возможных результата*: *успех* либо *неудача.*

Успех и неудачу при j-м испытании описывают события

$$У_j=\{u: u_j=1\},\ H_j=\{u: u_j=0\}\quad (j=1,\ldots,n),$$

составленные из строк, у которых на j-м месте соответственно 1 или 0. Верна следующая

Формула успехов и неудач:

$$P\Big(\bigcap_{i\in K}У_i\bigcap_{j\in L}H_j\Big)=a^k(1-a)^l.$$

В этой формуле K и L обозначают *непересекающиеся* множества, составленные соответственно из k и l номеров $1,\ldots,n$. В частности,

$$P(У_j)=a,\ P(H_j)=1-a\quad (j=1,\ldots,n).$$

Пример 1. *При рождении $n=2$ неидентичных близнецов каждый из них с вероятностью 0,52 оказывается мальчиком и с вероятностью $1-a=0{,}48$ — девочкой. Какова вероятность того, что близнецы оказываются одинакового пола?*

Решение. Предположим, что для описания пола неидентичных близнецов можно использовать модель Бернулли $n=2$ испытаний с вероятностью успеха $a=0{,}52$. В этой модели задача сводится к вычислению вероятности $P(A)$ события

$$A=У_1У_2+H_1H_2,$$

описывающего появление двух мальчиков или двух девочек. Используя правило сложения и формулу успехов и неудач, получаем:

$$P(A)=P(У_1У_2)+P(H_1H_2)=a^2+(1-a)^2=(0{,}52)^2+(0{,}48)^2\approx 0{,}51.$$

Пример 2 (рис. 35). *Нажимая кнопку K, можно привести в действие устройство P, если и только если замкнуты одинаковые и независимые блокировки 1 и 2, или 3, или 4. Каждая из этих блокировок с вероятностью $a=10^{-7}$ замыкается случайно. Какова вероятность случайной разблокировки системы?*

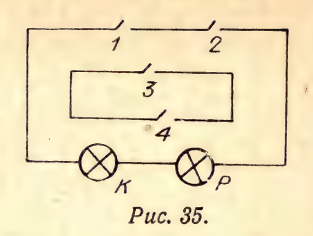


Рис. 35.

Решение. Используем модель Бернулли $n=4$ испытаний с вероятностью успеха $a=10^{-7}$. Задача сводится к вычислению вероятности объединения $A \cup B \cup C$ событий

$$A = У_1 У_2, \; B = У_3, \; C = У_4,$$

описывающих замыкание блокировок 1 и 2, 3, 4 соответственно. Используя правило объединения и формулу успехов и неудач, получаем:

$$P(B \cup C) = P(B) + P(C) - P(BC) = 2a - a^2.$$

Заметим, что $A(B \cup C) = \{1111, 1110, 1101\}$.

$$P(A(B \cup C)) = a^4 + 2a^3(1-a).$$

Еще раз используя *правило объединения*, находим:

$$P(A \cup B \cup C) = P(A) + P(B \cup C) - P(A(B \cup C)) =$$

$$= a^2 + 2a - a^2 - a^4 - 2a^3(1-a) = 2a - 2a^3 + a^4 < 10^{-6}.$$

Можно считать, что случайная разблокировка системы *практически невозможна.*

Пример 3. *На реактивном двигателе вместо одного установлено два воспламенителя 1 и 2. На воспламенителе 1 вместо одного установлено два электрозапала 3 и 4, на воспламенителе 2 — электрозапалы 5 и 6. Каждый из всех этих элементов системы запуска двигателя, независимо от других, с вероятностью $a=10^{-3}$ выходит из строя. Какова вероятность запуска двигателя?*

Решение. Используем модель Бернулли $n=6$ испытаний с вероятностью успеха $a=10^{-3}$. Задача сводится к вычислению вероятности объединения $A \cup B \cup C \cup D$ событий:

$$A = H_1 H_3, \; B = H_1 H_4, \; C = H_2 H_5, \; D = H_2 H_6,$$

описывающих рабочее состояние воспламенителей и соответствующих электрозапалов.

Последовательно применяя правило объединения, убеждаемся в том, что:

$$P(A \cup B \cup C \cup D) = S_1 - S_2 + S_3 - S_4,$$

где

$$S_1=P(A)+P(B)+P(C)+P(D),$$
$$S_2=P(AB)+P(AC)+P(AD)+P(BC)+P(BD)+P(CD),$$
$$S_3=P(ABC)+P(ABD)+P(ACD)+P(BCD),$$
$$S_4=P(ABCD).$$

Используя формулу успехов и неудач, получаем:

$$P(A\cup B\cup C\cup D)=4(1-a)^2-(2(1-a)^3+4(1-a)^4)+$$
$$+4(1-a)^5-(1-a)^6=1-a^2-2a^3+a^4+2a^5-a^6\approx 1-10^{-6}.$$

Можно считать, что запуск двигателя *практически достоверен.*

1.2.1. Задача о первом успехе

Какова вероятность того, что в последовательности n одинаковых и независимых испытаний, каждое из которых с вероятностью a оканчивается успехом и с вероятностью $1-a$ — неудачей, первый успех появляется при k-м испытании?

Решение. В модели Бернулли n испытаний с вероятностью успеха a задача сводится к вычислению вероятности события $A=У_k\bigcap\limits_{j<k}H_j$, описывающего успех при k-м и неудачи при всех предыдущих испытаниях $(1\leqslant k\leqslant n)$. По формуле успехов и неудач

$$P(A)=a(1-a)^{k-1}.$$

Пример 1. *Симметричная монета подбрасывается $n=10$ раз. Какова вероятность того, что первый герб появится при $k=10$-м подбрасывании?*

Ответ: $(1-1/2)^{10-1}\cdot 1/2=(1/2)^{10}\approx 0{,}001$.

Пример 2. *В одинаковых и независимых условиях производятся $n=3$ выстрела, при каждом из которых с вероятностью $a=0{,}8$ поражается цель, а с вероятностью $1-a=0{,}2$ — нет. Какова вероятность того, что цель поражается впервые при $k=3$-м выстреле?*

Ответ: $0{,}2^2\cdot 0{,}8=0{,}032$.

Пример 3. *Из 10 000 изделий, среди которых 10 негодных, $n=100$ раз наугад выбирается одно (и каждый раз возвращается на место.) Какова вероятность того, что первое негодное изделие появится при $k=10$-м выборе?*

Ответ: $(1-0{,}001)^9\cdot 0{,}001<0{,}001$.

2.2.1. Задача о хотя бы одном успехе

Какова вероятность того, что оканчивается успехом хотя бы одно из n одинаковых и независимых испытаний, каждое из которых с вероятностью a оканчивается успехом, а с вероятностью $1-a$ — неудачей?

Решение. В модели Бернулли n испытаний с вероятностью успеха a задача сводится к вычислению вероятности $P(A)$ события A, составленного из всех строк u, в которых есть хотя бы одна единица. Дополнение A' события A состоит из единственной строки $u=0\dots0$. Поэтому

$$P(A')=p(0\dots0)=a^0(1-a)^{n-0}=(1-a)^n.$$

По *правилу дополнения* получаем:

$$P(A)=1-(1-a)^n.$$

Пример 1. *Симметричная монета подбрасывается $n=10$ раз. Какова вероятность того, что хотя бы один раз она падает гербом вверх?*

Ответ: $1-(1/2)^{10}\approx0{,}999$.

Пример 2. *В одинаковых и независимых условиях производятся $n=3$ выстрела, при каждом из которых с вероятностью $a=0{,}8$ поражается цель, а с вероятностью $1-a=0{,}2$ — нет. Какова вероятность того, что цель поражается хотя бы один раз?*

Ответ: $1-(0{,}2)^3=0{,}992$.

Пример 3. *Из 10 000 изделий, среди которых 10 негодных, $n=100$ раз выбирается наугад одно (и каждый раз возвращается на место). Какова вероятность того, что среди выбираемых изделий хотя бы один раз оказывается негодное?*

Ответ: $1-(1-0{,}001)^{100}\approx0{,}1$.

3.2.1. Задача о числе испытаний

Каким должно быть число n одинаковых и независимых испытаний, каждое из которых с вероятностью a оканчивается успехом, а с вероятностью $1-a$ — неудачей, чтобы вероятность хотя бы одному испытанию окончиться успехом была больше $1-\alpha$?

Решение. Используя решение задачи о хотя бы одном успехе, находим, что требуемое в условии задачи неравенство

$$P(A)\geqslant1-\alpha$$

эквивалентно неравенству

$$(1-a)^n\leqslant\alpha.$$

Если $0<a<1$ и $0<\alpha<1$, то это неравенство эквивалентно неравенству

$$n\geqslant\frac{\log\alpha}{\log(1-a)}.$$

Пример 1. *Какое число n раз нужно подбросить симметричную монету, чтобы вероятность хотя бы одного появления герба была больше $1-\alpha=0{,}99$?*

Ответ:

$$n\geqslant\frac{\log(0{,}01)}{\log(1/2)}>6.$$

Замечание. Этот результат можно истолковать так: если подбрасывать симметричную монету 7 раз, то практически достоверно, что хотя бы один раз она упадет гербом вверх.

Упражнение. Подбросить монету 100 раз по 7 раз и отметить число семерок, в которых герб появится хотя бы один раз.

Пример 2. *При каждом выстреле с вероятностью $a=0,7$ поражается цель, а с вероятностью $1-a=0,3$ — нет. Какое число n выстрелов нужно произвести для того, чтобы вероятность хотя бы одного поражения цели была больше $1-\alpha=0,99$?*

Ответ:

$$n \geqslant \frac{\log(0,01)}{\log(0,3)} > 3.$$

Замечание. Этот результат можно истолковать так: если производить 4 выстрела, то практически достоверно, что хотя бы один раз цель окажется пораженной.

Пример 3. *Какое число n раз нужно наугад выбрать одно из 10 000 изделий, среди которых 10 негодных (каждый раз возвращая его на место), чтобы вероятность хотя бы один раз выбрать негодное изделие была больше $1-\alpha=0,999$?*

Ответ:

$$n \geqslant \frac{\log(0,001)}{\log(0,099)} \approx 7500.$$

Замечание. Этот результат объясняется малостью доли $a=0,001$ негодных изделий и чрезвычайной строгостью $1-\alpha=0,999$ контроля. Для того, чтобы обеспечить такой строгий контроль, нужно проверить примерно 3/4 всех изделий.

4.2.1. Задача о данном числе успехов

Какова вероятность того, что оканчиваются успехом ровно m из n одинаковых и независимых испытаний, каждое из которых с вероятностью a оказывается успехом, а с вероятностью $1-a$ — неудачей?

Решение. *В модели Бернулли n испытаний с вероятностью успеха a* задача сводится к вычислению вероятности $P(A_m)$ события

$$A_m = \{u : s(u) = m\},$$

составленного из всех строк $u = u_1 \ldots u_n$, в которых ровно m единиц. Число таких строк равно $\binom{n}{m}$. Элементарная вероятность $p(u)$ каждой такой строки u одна и та же:

$$p(u) = a^m (1-a)^{n-m}.$$

Поэтому

$$P(A_m) = \binom{n}{m} a^m (1-a)^{n-m}.$$

Пример 1. *Вычислительная машина производит $n=1\,000\,000$ одинаковых и независимых операций, в каждой из которых с вероятностью $a=0{,}001$ происходит ошибка, а с вероятностью $1-a=0{,}999$ — нет. Какова вероятность того, что все $n=1\,000\,000$ операций производятся машиной без ошибок?*

Ответ: $P(A_0)=(1-a)^n=(1-0{,}001)^{1\,000\,000}$.

Замечание. Неравенство Бернулли позволяет оценить малость этой вероятности.

Если $0<a<1$, то $c=\frac{1}{1-a}>1$, $c-1=\frac{a}{1-a}$,

$$(1-a)^n=\frac{1}{c^n}<\frac{1}{1+n(c-1)}<\frac{1}{n(c-1)}=\frac{1-a}{na}<\frac{1}{na}.$$

В частности, если $a=10^{-3}$ и $n=10^6$, то

$$(1-10^{-3})^{10^6}<\frac{1}{10^6 10^{-3}}=\frac{1}{10^3}=10^{-3}.$$

Пример 2. *Елочная гирлянда состоит из $n=10$ последовательно соединенных лампочек, каждая из которых с вероятностью $a=0{,}01$ перегорает, а с вероятностью $1-a=0{,}99$ — нет. Какова вероятность того, что ни одна лампочка в гирлянде не перегорает?*

Ответ: $P(A_0)=(1-a)^n=(1-10^{-2})^{10}\approx 1-10\cdot 10^{-2}=0{,}9$.

Замечание. Это число оценивает *надежность* гирлянды лампочек.

Пример 3. *Имеются две розовые урны, в каждой из которых находятся красный и белый шары. Из каждой урны вынимается наугад выбранный шар. Какова вероятность того, что среди них $k=0, 1, 2$ красных?*

Решение. Условие о выборе наугад позволяет использовать модель Бернулли $n=2$ испытаний с вероятностью успеха $a=1/2$. Задача сводится к вычислению вероятностей событий A_0, A_1, A_2:

$$P(A_0)=\binom{2}{0}(1/2)^0(1-1/2)^{2-0}=1/4,$$

$$P(A_1)=\binom{2}{1}(1/2)^1(1-1/2)^{2-1}=1/2,$$

$$P(A_2)=\binom{2}{2}(1-1/2)^{2-2}(1/2)^2=1/4.$$

Замечание. Задача о розовых урнах имеет отношение к генетике. С основами генетики можно познакомиться, прочитав книгу Н. П. Дубинина «Горизонты генетики». (М., «Просвещение», 1970).

При скрещивании растений двух родственных видов, одного с белыми цветами, а другого с красными, были получены гибриды с розовыми цветами. При скрещивании 564 таких гибридов 141 из них дали растение с белыми цветами, 291 — с розовыми и 132 —

с красными. Соответствующие частоты 0,350; 0,516 и 0,234 близки к 1/4, 1/2 и 1/4.

Теория Менделя объясняет этот факт следующим образом. Каждая участвующая в образовании *гамет* клетка растения содержит пару *хромосом,* у которой связанный с окраской цветов *локус* занят парой *генов* u_1, u_2. Принадлежащий 1-й хромосоме ген u_1, может быть геном 0, обусловливающим белую окраску цветов, либо геном 1, обусловливающим красную окраску. Аналогично определяется принадлежащий 2-й хромосоме ген u_2. Если растение имеет *генотип* 00 ($u_1=u_2=0$), то у него белые цветы, если 11 ($u_1=$ $=u_2=1$),— красные. *Гибриды* имеют генотип {01, 10} ($u_1=0$, $u_2=1$ или $u_1=1$, $u_2=0$) и розовые цветы. Рассматриваемые клетки гибридов при делении производят гаметы двух типов: одна из них содержит хромосому с геном 0, а другая — хромосому с геном 1. При скрещивании гаметы одной клетки соединяются с гаметами другой независимым и одинаково случайным образом: каждая гамета каждого типа с одинаковой вероятностью соединяется с каждой другой.

Схематически этот процесс можно описать выбором наугад шаров из урн, рассматривавшимся в примере 3.

Пример 4. *Каково наиболее вероятное число успехов для одинаковых и независимых испытаний, каждое из которых оканчивается с вероятностью a успехом, а с вероятностью 1—a — неудачей?*

Решение. В модели Бернулли n испытаний с вероятностью успеха a задача сводится к выяснению того, какое из чисел $P(A_m)$ наибольшее ($m=0, 1, \ldots, n$). Будем предполагать, что $0<a<1$.

Как нетрудно проверить, использовав решение задачи о данном числе успехов и свойства биномиальных коэффициентов,

$$\frac{P(A_m)}{P(A_{m-1})}=\frac{(n+1)a-m}{m(1-a)}+1.$$

Поэтому

$$P(A_{m-1})>P(A_m), \text{ если } m>(n+1)a,$$
$$P(A_{m-1})=P(A_m), \text{ если } m=(n+1)a,$$
$$P(A_{m-1})<P(A_m), \text{ если } m<(n+1)a.$$

Следовательно, наиболее вероятное число успехов $\hat{m}$ равно целой части числа $(n+1)a$: $\hat{m}=[(n+1)a]$.

Замечание. Если число $(n+1)a$ *целое*, то наиболее вероятными числами успехов являются $\hat{m}=(n+1)a$, $\hat{m}-1=(n+1)a-1$. Например, если $n=3$ и $a=1/2$, то

$$P(A_0)=1/8,\ P(A_1)=3/8,\ P(A_2)=3/8,\ P(A_3)=1/8.$$

Пример 5. *Какова вероятность того, что при 100 подбрасываниях симметричной монеты ровно 50 раз появится герб?*

Решение. В модели Бернулли $n=100$ испытаний с вероятностью успеха $a=1/2$ задача сводится к вычислению вероятности $P(A_{50})$ события A_{50}:

$$P(A_{50}) = \binom{100}{50} (1/2)^{50} (1-1/2)^{100-50} = \frac{100!}{50!\,50!} 2^{-100} < 0{,}08.$$

Заметим, что при $n=100$ и $a=1/2$ число

$$\hat{m}=[(n+1)a]=[101\cdot 1/2]=[50+1/2]=50$$

является наиболее вероятным числом успехов.

5.2.1. Задача о большом числе успехов

Какова вероятность того, что оканчиваются успехом больше, чем m из n одинаковых испытаний, каждое из которых оканчивается с вероятностью a успехом, а с вероятностью $1-a$ — неудачей?

Решение. В модели Бернулли n испытаний с вероятностью успеха a задача сводится к вычислению вероятности $p(A)$ суммы

$$A=A_m+\ldots+A_n \quad (0\leqslant m\leqslant n)$$

попарно не пересекающихся событий $A_m, \ldots, A_n$.

Используя правило сложения и решение задачи о данном числе успехов, получаем:

$$P(A) = P(A_m) + \ldots + P(A_n) =$$

$$= \binom{n}{m} a^m (1-a)^{n-m} + \ldots + \binom{n}{n} a^n (1-a)^{n-n} =$$

$$= \sum_{m\leqslant l\leqslant n} \binom{n}{l} a^l (1-a)^{n-l}.$$

Замечание. При большом числе слагаемых использовать полученное равенство трудно. Если $m\geqslant na$, то можно оценить найденную сумму следующим образом. Будем предполагать, что $0<a<1$.

Как уже отмечалось,

$$\frac{P(A_l)}{P(A_{l-1})} = \frac{(n+1)a-l}{l(1-a)} + 1 = q(l).$$

Если $(n+1)a\leqslant m+1\leqslant l$, то

$$q(l-1)\geqslant q(l),$$

и поэтому

$$q=q(m+1)\geqslant q(l).$$

Следовательно,

$$\frac{P(A_l)}{P(A_m)} = \frac{P(A_l)}{P(A_{l-1})} \frac{P(A_{l-1})}{P(A_{l-2})} \cdots \frac{P(A_{m+1})}{P(A_m)} =$$

$$=q(l)\,q(l-1)\,\ldots\,q(m+1)\leqslant q^{l-m}.$$

Таким образом,

$$P(A_l) \leqslant P(A_m) q^{l-m} \quad (m \leqslant l \leqslant n).$$

Суммируя эти неравенства, получаем:

$$\sum_{m \leqslant l \leqslant n} P(A_l) \leqslant P(A_m) \sum_{m \leqslant l \leqslant n} q^{l-m}.$$

Заметим, что

$$0 < q = q(m+1) = \frac{(n+1)a - (m+1)}{(m+1)(1-a)} + 1 < 1.$$

Используя известное равенство для суммы геометрической прогрессии, находим:

$$\sum_{m \leqslant l \leqslant n} q^{l-m} = \sum_{0 \leqslant k \leqslant n-m} q^k < \sum_{0 \leqslant k} q^k = \frac{1}{1-q} = \frac{(m+1)(1-a)}{(m+1)-(n+1)a}.$$

Следовательно,

$$P(A) = \sum_{m \leqslant l \leqslant n} P(A_l) < P(A_m) \frac{(m+1)(1-a)}{(m+1)-(n+1)a}.$$

Пример 1. *Производится залп $n=12$ одинаковых и независимых ракет, каждая из которых с вероятностью $a=1/3$ поражает цель, а с вероятностью $1-a=2/3$ — нет. Цель уничтожается, если ее поражают больше $m=8$ ракет. Какова вероятность уничтожения цели?*

Решение. Используя решение задачи о большом числе успехов и полученную оценку, находим, что искомая вероятность

$$P(A) = \sum_{8 \leqslant l \leqslant 12} \binom{12}{l} \left(\frac{1}{3}\right)^l \left(\frac{2}{3}\right)^{12-l} < \binom{12}{8} \left(\frac{1}{3}\right)^8 \left(\frac{2}{3}\right)^4 \frac{2(8+1)\,2/3}{(8+1)-(12+1)\,1/3} < 0{,}02.$$

Из-за малой вероятности поражения и большого количества ракет, необходимых для ее уничтожения, цель практически не уничтожается.

Пример 2. *Космическая частица, попадая в данную область пространства, порождает лавину $n=600$ одинаковых и независимых частиц, каждая из которых с вероятностью $a=1/2$ регистрируется одним из счетчиков, а с вероятностью $1-a=1/2$ — нет. Какова вероятность того, что регистрируется больше, чем $m=500$ частиц?*

Решение. Используя решение задачи о большом числе успехов и полученную оценку, находим, что искомая вероятность

$$P(A) = \sum_{l=500}^{600} \binom{600}{l} (1/2)^l (1-1/2)^{600-l} \leqslant \binom{600}{500} 2^{-600} \frac{501 \cdot 1/2}{501 - 600 \cdot 1/2} \leqslant$$

$$\leqslant \frac{600 \cdot 599 \ldots 501}{100 \cdot 99 \ldots 1} \cdot \frac{501}{1002-601} 2^{-600} \leqslant 2 \cdot 10^{-10}.$$

Регистрация счетчиками такого большого числа частиц в данных условиях практически невозможна.

Пример 3. *Что более вероятно получить*:

1) *хотя бы 1 раз 6 очков, подбрасывая кость 6 раз*;
2) *хотя бы 2 раза 6 очков, подбрасывая кость 12 раз*;
3) *хотя бы 3 раза 6 очков, подбрасывая кость 18 раз?*

Решение. Используем модель Бернулли $n=6, 12, 18$ испытаний с вероятностью успеха $a=1/6$. Задача сводится к вычислению вероятностей событий $B_{mn}=A_m+\ldots+A_n$, описывающих появление больше $m=1, 2, 3$ успехов соответственно:

$$P(B_{mn})=\sum_{m\leqslant l\leqslant n}\binom{n}{l}(1/6)^l(5/6)^{n-l}=1-\sum_{0<k<m}\binom{n}{k}(1/6)^k(5/6)^{n-k}.$$

Произведя вычисления, получаем:

1) $P(B_{1,6})=1-(5/6)^6\approx 0{,}665$;

2) $P(B_{2,12})=(1-5/6)^{12}-12\cdot 1/6(5/6)^{11}\approx 0{,}619$;

3) $P(B_{3,18})=1-(5/6)^{18}-18\cdot 1/6(5/6)^{17}-153(1/6)^2(5/6)^{16}\approx 0{,}597$.

Замечание. Вместо того, чтобы подбрасывать $n=6, 12, 18$ раз одну кость, можно один раз подбросить $n=6, 12, 18$ костей. Связанные с такими подбрасываниями игры разработал Сэмуэль Пепайс. Вероятности появления соответственно хотя бы $m=1, 2, 3$ шестерок по просьбе Пепайса вычислил Исаак Ньютон.

§ 2. УСЛОВНАЯ ВЕРОЯТНОСТЬ

Для каждых события A и *возможного* события B, для которого $P(B)>0$, определяется *условная* вероятность

$$P_B(A)=P(AB)/P(B)$$

события A при условии B. Условная вероятность $P_B(A)$ служит *мерой реализуемости* события *A при условии*, что реализуется событие B.

1.2. Правила умножения и деления

Определение условной вероятности можно сформулировать как

Правило деления: $P_B(A)=P(AB)/P(B)$.

Непосредственно из правила деления вытекает

Правило умножения: $P(AB)=P(B)P_B(A)$.

Задача 1. *Игральная кость подбрасывается два раза. Известно, что сумма очков равна 10. Какова вероятность при этом условии того, что один раз появляется 6 очков?*

Решение. Используем модель задачи 1 пункта 1 из § 1. Задача сводится к вычислению условной вероятности $P_B(A)$ события

$A=\{46, 64\}$ при условии $B=\{46, 55, 64\}$. По правилу деления $P_B(A)=P(AB)/P(B)=P(\{46, 64\})/P(\{46, 55, 64\})=$
$$=(2/36)/(3/36)=2/3.$$

Задача 2. *Рассматривается выбор с возвращением, описанный в задаче 2 пункта 1 из § 1. Известно, что первый раз выбирается шар 1. Какова вероятность при этом условии того, что второй раз выбирается шар 2?*

Решение. В модели задачи 2 пункта 1 из § 1 дело сводится к вычислению условной вероятности $P_B(A)$ события $A=\{12, 22, 32, 42, 52\}$ при условии $B=\{11, 12, 13, 14, 15\}$. По правилу деления,
$$P_B(A)=P(AB)/P(B)=P(\{12\})/P(\{11, 12, 13, 14, 15\})=1/5.$$

Задача 3. *Рассматривается выбор без возвращения, описанный в задаче 3 пункта 1 из § 1. Известно, что первый раз выбирается шар 1. Какова вероятность при этом условии того, что второй раз выбирается шар 2?*

Решение. В модели задачи 3 пункта 1 из § 1 дело сводится к вычислению условной вероятности $P_B(A)$ события $A=\{12, 32, 42, 52\}$ при условии $B=\{12, 13, 14, 15\}$. По правилу деления,
$$P_B(A)=P(AB)/P(B)=P(\{12\})/P(\{12, 13, 14, 15\}=1/4.$$

Задача 4. *Симметричная монета подбрасывается $n=10$ раз. Известно, что при $k=3$-м подбрасывании появляется герб. Какова вероятность при этом условии того, что этот герб первый?*

Решение. В модели Бернулли $n=10$ испытаний с вероятностью успеха $a=1/2$ задача сводится к вычислению условной вероятности $P_B(A)$ события $A=H_1H_2У_3$ при условии $B=У_3$. По правилу деления и формуле прямоугольника,
$$P_B(A)=P(AB)/P(B)=P(H_1H_2У_3)/P(У_3)=1/4.$$

Задача 5. *Рассматривается выбор без возвращения, описанный в задаче 3 пункта 1 из § 1. Какова вероятность того, что первый раз выбирается шар 1, а второй раз — шар 2?*

Решение. В модели задачи 3 пункта 1 из § 1 дело сводится к вычислению вероятности $P(AB)$ пересечения событий $A=\{12, 32, 42, 52\}$ и $B=\{12, 13, 14, 15\}$. Используя правило умножения и результат задачи 3 этого пункта, получаем:
$$P(AB)=P(B)P_B(A)=(4/20)(1/4)=1/20.$$

То же самое получается и при непосредственном расчете:
$$P(AB)=P(\{12\})=1/20.$$

2.2. Формула полной вероятности

Стандартные задачи, решаемые с помощью формулы полной вероятности, можно схематически описать следующим образом.

Задача о полной вероятности. *Известны:*

1) вероятности $P(B_i)=\beta_i$ нескольких исключающих друг друга условий B_i, одно из которых с достоверностью выполняется;

2) условные вероятности $P_{B_i}(A)=\alpha_i$ события A при условии, что выполняется B_i.

Какова вероятность $P(A)$ события A?

Решение. Рассмотрим конечную вероятностную модель, определяемую множеством исходов $U=\{11, \ldots, i1, \ldots, n1, 10, \ldots, i0, \ldots, n0\}$, составленным из n строк $11, \ldots, i1, \ldots, n1$, n строк $10, \ldots, i0, \ldots, n0$, и элементарной вероятностью p со значениями

$$p(i1)=\beta_i\alpha_i,\ p(i0)=\beta_i(1-\alpha_i)\quad (i=1, \ldots, n).$$

В этой модели каждое условие

$$B_i=\{i1, i0\}\quad (i=1, \ldots n)$$

составлено из двух строк $i1$ и $i0$, а событие

$$A=\{11, \ldots, i1, \ldots, n1\}$$

— из n строк $11, \ldots, i1, \ldots, n1$.

Используя свойства вероятностей β_i и α_i, нетрудно проверить, что в этой модели

$$P(B_i)=\beta_i,\quad P_{B_i}(A)=\alpha_i\quad (i=1, \ldots, n).$$

Решение задачи дает

Формула полной вероятности:

$$P(A)=\sum_{i=1}^{n} P(B_i)\,P_{B_i}(A)=\sum_{i=1}^{n}\beta_i\alpha_i.$$

Пример 1. *Имеются красная, черная и белая урны с черными и белыми шарами.*

В красной урне $a_1=1$ черный и $a_2=2$ белых шара,
в черной — $b_1=2$ черных и $b_2=3$ белых,
в белой — $c_1=3$ черных и $c_2=4$ белых.

Из красной урны наугад выбирается шар. Если этот шар черный, то следующий шар также наугад выбирается из черной урны. Если шар, выбранный из красной урны, белый, то следующий шар наугад выбирается из белой урны. Какова вероятность того, что оба выбирающиеся шара имеют одинаковый цвет?

1. **Подробное решение.** Условимся описывать:

строкой 11 выбор *черного* шара из *красной* урны и
строкой 11 — *черного* шара из *черной* урны;
строкой 10 выбор *черного* шара из *красной* урны и
строкой 10 — *белого* шара из *черной* урны;
строкой 21 выбор *белого* шара из *красной* урны и
строкой 21 — *белого* шара из *белой* урны;

строкой 20 выбор *белого* шара из *красной* урны и строкой 20 — *черного* шара из *белой* урны.

Множество строк

$$U=\{11, 21, 10, 20\}$$

описывает возможные результаты рассматриваемого опыта.

Условия

$$B_1=\{11, 10\},\ B_2=\{21, 20\}$$

описывают соответственно выбор черного и белого шаров из красной урны. Так как выбор производится *наугад*, то

$$P(B_1)=\frac{a_1}{a_1+a_2}=1/3,\quad P(B_2)=\frac{a_2}{a_1+a_2}=2/3.$$

Аналогично, вследствие выбора наугад шаров из черной и белой урн, условные вероятности события

$$A=\{11, 21\},$$

описывающего выбор двух черных шаров или двух белых, равны соответственно:

$$P_{B_1}(A)=\frac{b_1}{b_1+b_2}=2/5,\quad P_{B_2}(A)=\frac{c_2}{c_1+c_2}=4/7.$$

Из правила умножения вытекает, что элементарная вероятность p, описывающая реализуемость возможных результатов рассматриваемого опыта, имеет значения:

$$p=(11)\frac{a_1}{a_1+a_2}\cdot\frac{b_1}{b_1+b_2}=\frac{1}{3}\cdot\frac{2}{5},\quad p(21)=\frac{a_2}{a_1+a_2}\cdot\frac{c_2}{c_1+c_2}=\frac{2}{3}\cdot\frac{4}{7},$$

$$p(10)=\frac{a_1}{a_1+a_2}\cdot\frac{b_2}{b_1+b_2}=\frac{1}{3}\cdot\frac{3}{5},\quad p(20)=\frac{a_2}{a_1+a_2}\cdot\frac{c_1}{c_1+c_2}=\frac{2}{3}\cdot\frac{3}{7}.$$

В вероятностной модели, определяемой такими множеством исходов U и элементарной вероятностью p, задача сводится к вычислению вероятности $p(A)$ события $A=\{11, 21\}$. По формуле полной вероятности

$$P(A)=P(B_1)P_{B_1}(A)+P(B_2)P_{B_2}(A)=\frac{1}{3}\cdot\frac{2}{5}+\frac{2}{3}\cdot\frac{4}{7}\approx 0{,}51.$$

2. **Краткое решение.** Задача сводится к задаче о полной вероятности с вероятностями условий $\beta=1/3$, $\beta_2=2/3$ и условными вероятностями $\alpha_1=2/5$, $\alpha_2=4/7$. Искомая вероятность равна

$$\beta_1\alpha_1+\beta_2\alpha_2\approx 0{,}51.$$

Пример 2. *По данным переписи 1951 года, в Англии и Уэльсе среди отцов, имеющих сыновей, оказалось*

13% темноглазых и 87% светлоглазых.

У темноглазых отцов оказалось

39% темноглазых и 61% светлоглазых сыновей.

У светлоглазых отцов оказалось

10% темноглазых и 90% светлоглазых сыновей.

Какова вероятность того, что наугад выбранные среди этого населения отец и сын имеют глаза одинакового цвета?

Решение. Задача сводится к задаче о полной вероятности с вероятностями условий $\beta_1=0{,}13$, $\beta_2=0{,}87$ и условными вероятностями $\alpha_1=0{,}39$, $\alpha_2=0{,}90$. Искомая вероятность равна

$$\beta_1\alpha_1+\beta_2\alpha_2\approx 0{,}82.$$

Замечание. Выбор наугад отца и сына можно представлять как выбор двух шаров по схеме примера 1 из красной, черной и белой урн с $a_1=13$ и $a_2=87$, $b_1=39$ и $b_2=61$, $c_1=10$ и $c_2=90$ черными и белыми шарами соответственно. Черные и белые шары в красной урне описывают темноглазых отцов, в черной — темноглазых и светлоглазых сыновей у темноглазых отцов, в белой — у светлоглазых отцов.

Пример 3. *Статистика показывает, что среди двоен оказывается*

28% идентичных и 72% неидентичных близнецов.

Среди идентичных близнецов

100% одного пола, 0% разного пола.

Среди неидентичных близнецов

50% одного пола, 50% разного пола.

Какова вероятность того, что наугад выбранные среди двоен близнецы имеют одинаковый пол?

Решение. Задача сводится к задаче о полной вероятности с вероятностями условий $\beta_1=0{,}28$, $\beta_2=0{,}72$ и условными вероятностями $\alpha_1=1$, $\alpha_2=0{,}5$. Искомая вероятность равна

$$\beta_1\alpha_1+\beta_2\alpha_2\approx 0{,}64.$$

Замечание. Эта вероятность близка наблюдаемой в действительности частоте появления однополых близнецов при рождении двоен.

Появление однополых близнецов можно представить как появление двух шаров одинакового цвета по схеме примера 1 из красной, черной и белой урн с $a_1=28$ и $a_2=72$, $b_1=100$ и $b_2=0$, $c_1=50$ и $c_2=50$ черными и белыми шарами соответственно. Черные и белые шары в красной урне описывают идентичных и неидентичных близнецов, в черной — однополых и разнополых близнецов среди идентичных, в белой — среди неидентичных.

Пример 4. *В условиях примера 1 какова вероятность того, что второй из выбирающихся шаров черный?*

Решение. Задача сводится к задаче о полной вероятности с вероятностями условий $\beta_1=1/3$, $\beta_2=2/3$ и условными вероятностями $\alpha_1=2/5$, $\alpha_2=3/7$. Искомая вероятность равна

$$\beta_1\alpha_1+\beta_2\alpha_2\approx 0{,}42.$$

Замечание. Вероятностная модель для примера 4 аналогична модели примера 1. Меняется только содержание исходов 21 и 20: в модели для примера 4 исход 21 описывает выбор белого шара из

красной урны и черного из белой, а исход 20 — белого из красной и белого из белой. В соответствии с этим меняются элементарные вероятности исходов и условная вероятность $P_{B_2}(A)$ события A:

$$P_{B_2}(A) = \frac{c_1}{c_1 + c_2} = 3/7.$$

Следующие примеры 5, 6 решаются по той же схеме, что и пример 4.

Пример 5. *Среди помещенных в Т-образный лабиринт голодных крыс*

50% бегут в левый конец и 50% — в правый.

Среди крыс, побывавших в левом конце с пищей и вновь помещенных в лабиринт,

$(50+100\cdot\varepsilon)$% бегут в левый конец и $(50-100\cdot\varepsilon)$% — в правый.

Среди крыс, побывавших в правом конце без пищи и вновь помещенных в лабиринт,

50% бегут в левый конец и 50% — в правый.

Какова вероятность того, что вновь помещенная в лабиринт крыса побежит в левый конец?

Решение. Задача сводится к задаче о полной вероятности с вероятностями условий $\beta_1=0{,}5$, $\beta_2=0{,}5$ и условными вероятностями $\alpha_1=0{,}5+\varepsilon$, $\alpha_2=0{,}5$ $(0\leqslant\varepsilon\leqslant0{,}5)$. Искомая вероятность равна

$$\beta_1\alpha_1+\beta_2\alpha_2=0{,}5\ (1+\varepsilon).$$

Замечание. Число ε выражает эффективность рассматриваемого процесса обучения крысы и определяется экспериментально. Например, если $\varepsilon=0{,}05$, то вероятность того, что повторно помещенная в лабиринт крыса побежит к пище, равна 0,525.

Пример 6. *В самоанском письменном тексте*

67% гласных и 33% согласных букв.

Среди букв, следующих непосредственно за гласной,

49% гласных и 51% согласных.

Среди букв, следующих непосредственно за согласной,

100% гласных и 0% согласных.

Какова вероятность того, что за наугад выбранной буквой самоанского текста непосредственно следует гласная буква?

Решение. Задача сводится к задаче о полной вероятности с вероятностями условий $\beta_1=0{,}67$, $\beta_2=0{,}33$ и условными вероятностями $\alpha_1=0{,}49$, $\alpha_2=1$. Искомая вероятность равна

$$\beta_1\alpha_1+\beta_2\alpha_2\approx0{,}66.$$

Замечание. В примерах 7—9 рассматривается система связи (рис. 36):

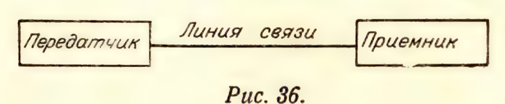


Рис. 36.

Передатчик посылает *сигнал* в *линию связи*. Проходя по линии связи, переданный сигнал может *исказиться* и превратиться в другой сигнал. *Приемник* улавливает сигнал, пришедший по линии связи. Таким образом, может быть принят не тот сигнал, который был передан.

В простейшем случае передатчик посылает сигналы 1 и 0 («—» и «+»). Проходя по линии связи, 1 может превратиться в 0 («—» может превратиться в «+») и наоборот. Таким образом, если передан сигнал 1 («—»), то может быть принят как сигнал 1 («—»), так и сигнал 0 («+»). Аналогично обстоит дело при передаче сигнала 0 («+»).

О системах связи и относящихся к ним интересных задачах можно прочитать в книге А. М. Яглома и И. М. Яглома «Вероятность и информация» (М., «Наука», 1973).

Пример 7. *По линии связи с вероятностями* $p_1=0{,}6$, $p_0=0{,}4$ *посылаются сигналы 1,0.*
Если посылается сигнал 1, то с вероятностями
$r_{11}=0{,}9$, $r_{10}=0{,}1$
принимаются сигналы 1,0.
Если посылается сигнал 0, то с вероятностями
$r_{01}=0{,}3$, $r_{00}=0{,}7$
принимаются сигналы 1,0.
Какова вероятность того, что принимается сигнал 1?

Решение. Задача сводится к задаче о полной вероятности с вероятностями условий $\beta_1=p_1$, $\beta_2=p_0$ и условными вероятностями

$$\alpha_1=r_{11},\ \alpha_2=r_{01}\ (p_1,\ p_0,\ r_{11},\ r_{10},\ r_{01},\ r_{00}\geqslant 0,\ p_1+p_0=r_{11}+r_{10}=$$

$$=r_{01}+r_{00}=1).$$

Искомая вероятность равна

$$\beta_1\alpha_1+\beta_2\alpha_2=p_1r_{11}+p_0r_{01}=q_1=0{,}66.$$

Пример 8. *В условиях примера 7 какова вероятность того, что принимается сигнал 0?*

Ответ: $p_1r_{10}+p_0r_{00}=q_0=0{,}34$.

Пример 9. *По линии связи с вероятностью* p_i *посылается сигнал* $i(i=0,\ \ldots,\ n-1)$. *Если посылается сигнал* i, *то с вероятностью* r_{ij} *принимается сигнал* j $(j=0,\ \ldots,\ n-1)$. *Какова вероятность того, что принимается сигнал* j?

Решение. Задача сводится к задаче о полной вероятности с вероятностями условий $\beta_i=p_i$ и условными вероятностями $\alpha_i=r_{ij}$. Искомая вероятность равна

$$q_j=\sum_{i=0}^{n-1}p_ir_{ij}\left(p_ir_{ij}\geqslant 0;\quad \sum_{i=0}^{n-1}p_i=1,\quad \sum_{j=0}^{n-1}r_{ij}=1\right).$$

Замечание. Если $r_{ij}=1$ при $i=j$ и $r_{ij}=0$ при $i\neq j$, то передача происходит без искажений и $q_i=p_j$ $(j=0,\ \ldots,\ n-1)$.

Если $r_{ij}=1/n$ для каждых i и j, то $q_j=1/n(j=0,\ \ldots,\ n-1)$. Прием описывается той же вероятностной моделью, что и выбор наугад.

3.2. Формула Байеса

Стандартные задачи, решаемые с помощью *формулы Байеса*, можно схематически описать следующим образом.

Задача о вероятностях гипотез. *Известны*:

1) вероятности $P(B_i) = \beta_i$ возможных исключающих друг друга предположений B_i, одно из которых верно;

2) условные вероятности $P_{B_i}(A) = \alpha_i$ события A при условии, что верно предположение B_i.

Какова условная вероятность $P_A(B_i)$ того, что верно предположение B_i при условии, что реализуется событие A?

Решение. Условия задачи о вероятностях гипотез являются перефразировкой условий задачи о полной вероятности. Поэтому для решения этих задач можно использовать одну и ту же вероятностную модель.

Решение задачи о вероятностях гипотез в этой модели дает **Формула Байеса:**

$$P_A(B_i) = \frac{P(B_i)\,P_{B_i}(A)}{\sum P(B_j)\,P_{B_j}(A)} = \frac{\beta_i\alpha_i}{\sum \beta_j\alpha_j}.$$

Пример 1. *В схеме примера 4 пункта 2 какова условная вероятность того, что первый шар черный при условии, что последний шар черный?*

Решение. В модели примера 4 пункта 1 задача сводится к вычислению условной вероятности $P_A(B_1)$ события $B_1 = \{11, 10\}$ при условии $A = \{11, 21\}$. По формуле Байеса

$$P_A(B_1) = \frac{P(B_1)\,P_{B_1}(A)}{P(B_1)\,P_{B_1}(A) + P(B_2)\,P_{B_2}(A)} \approx 0{,}32.$$

Тот же самый результат можно получить, используя вычисления примера 4 пункта 1 и правило деления:

$$P_A(B_1) = \frac{P(AB_1)}{P(A)} \approx 0{,}32.$$

Замечание. Следующие примеры 2 и 4 решаются по той же схеме, что и пример 1. Вопрос примера 3 аналогичен вопросу о вероятности вынуть первый белый шар при условии, что последний черный.

Пример 2. *В схеме примера 5 пункта 2 какова условная вероятность того, что крыса бегала к пище в первый раз при условии, что она побежала к пище во второй раз?*

Ответ: $\dfrac{\beta_1\alpha_1}{\beta_1\alpha_1 + \beta_2\alpha_2} = \dfrac{0{,}5\,(0{,}5+\varepsilon)}{0{,}5\,(0{,}5+\varepsilon) + 0{,}5\cdot 0{,}5} = \dfrac{0{,}5+\varepsilon}{1+\varepsilon}.$

В частности, если $\varepsilon = 0{,}05$, то искомая вероятность равна приблизительно 0,524.

Пример 3. *В схеме примера 6 пункта 2 какова условная вероятность того, что выбираемая буква оказывается согласной при условии, что непосредственно следующая за ней буква гласная?*

Ответ: $\dfrac{\beta_2\alpha_2}{\beta_1\alpha_1+\beta_2\alpha_2}=\dfrac{0{,}33\cdot 1}{0{,}67\cdot 0{,}49+0{,}33\cdot 1}\approx 0{,}5.$

Пример 4. *В схеме примера 7 пункта 2 какова условная вероятность того, что посылается сигнал 1 при условии, что принимается сигнал 1?*

Ответ: $\dfrac{\beta_1\alpha_1}{\beta_1\alpha_1+\beta_2\alpha_2}=\dfrac{p_1r_{11}}{p_1r_{11}+p_0r_{01}}=s_{11}=\dfrac{54}{66}\approx 0{,}82.$

Пример 5. *В схеме примера 8 пункта 2 какова условная вероятность того, что посылается сигнал 1 при условии, что принимается сигнал 0?*

Решение. Задача сводится к задаче о вероятностях гипотез с вероятностями предположений $\beta_i=p_i$ и условными вероятностями $\alpha_i=r_{ij}$. Искомая вероятность равна

$$s_{ij}=\frac{p_ir_{ij}}{\sum\limits_{k=0}^{n-1}p_kr_{kj}}\quad(i=0,\ldots,n-1;\quad j=0,\ldots,n-1).$$

Замечание. Если $r_{ij}=1$ при $i=j$, $r_{ij}=0$ при $i\neq j$, то передача происходит без искажений и $s_{ij}=1$ при $j=i$ и $s_{ij}=0$ при $j\neq i$.

Если $r_{ij}=1/n$ для каждых i и j, то $s_{ij}=p_i$: условная вероятность посылаемого сигнала не зависит от принимаемого сигнала.

§ 3. НЕЗАВИСИМОСТЬ И ЗАВИСИМОСТЬ

События $C_1,\ldots,C_n$ называются *независимыми*, если вероятность пересечения каждых выбранных из них m событий $C_{i_1},\ldots,C_{i_m}$ равна *произведению* вероятностей этих событий:

$$P\left(C_{i_1},\ldots,C_{i_m}\right)=P\left(C_{i_1}\right)\ldots P\left(C_{i_m}\right).$$

Если для некоторых m событий $C_{i_1},\ldots,C_{i_m}$, выбранных из n событий $C_1,\ldots,C_n$, это равенство *не верно*, то события $C_1,\ldots,C_n$ называются *зависимыми* ($m\leqslant n$, $1\leqslant i_1,\ldots,i_m\leqslant n$).

В частности, при $n=2$ события $C_1=A$ и $C_2=B$ независимы, если и только если

$$P(AB)=P(A)P(B).$$

Если событие B *возможно* ($P(B)>0$), то это равенство эквивалентно равенству

$$P_B(A)=P(A).$$

В качестве *меры зависимости* событий A и B часто используется *коэффициент корреляции*

$$K(A, B) = \frac{P(AB) - P(A)P(B)}{\sqrt{P(A)(1 - P(A)P(B)(1 - P(B))}}$$

событий A и $B (0 < P(A),\ P(B) < 1)$. События A и B *независимы*, если и только если

$$K(A, B) = 0.$$

В связи с использованием коэффициента корреляции вместо *зависимость* событий говорят также *корреляция* событий.

Пример 1. В условиях примера 1 пункта 2 из § 2 события A и B_1, A и B_2 зависимы:

$$P_{B_1}(A) = \frac{2}{5} \neq \frac{54}{105} = P(A); \quad P_{B_2}(A) = \frac{4}{7} \neq \frac{54}{105} = P(A);$$

$$K(A, B_1) = \left(\frac{1}{3}\cdot\frac{2}{5} - \frac{1}{3}\cdot\frac{54}{105}\right)\Bigg|\sqrt{\frac{1}{3}\cdot\frac{2}{3}\cdot\frac{54}{105}\cdot\frac{51}{105}} = -\frac{2}{3\sqrt{17}} \approx -0{,}17;$$

$$K(A, B_2) = \frac{\frac{2}{3}\cdot\frac{4}{7} - \frac{2}{3}\cdot\frac{54}{105}}{\sqrt{\frac{1}{3}\cdot\frac{2}{3}\cdot\frac{54}{105}\cdot\frac{51}{105}}} = \frac{2}{3\sqrt{17}} \approx +0{,}17.$$

Замечание. Результат

$$K(A, B_2) = -K(A, B_1)$$

не случаен. Он объясняется тем, что событие B_2 является дополнением события B_1.

Пример 2. В условиях примера 4 пункта 2 из § 2 события A и B_1, A и B_2 *зависимы*:

$$P_{B_1}(A) = \frac{2}{5} \neq \frac{44}{105} = P(A); \quad P_{B_2}(A) = \frac{3}{7} \neq \frac{44}{105} = P(A);$$

$$K(A, B_1) \approx 0{,}03; \ K(A, B_2) \approx 0{,}03.$$

Замечание. Если состав шаров в черной и белой урнах одинаков ($b_1 = c_1$, $b_2 = c_2$), то в примере 2 события A и B_1, A и B_2 *независимы*:

$$P_{B_1}(A) = P_{B_2}(A) = P(A) = b_1/(b_1 + b_2) = c_1/(c_1 + c_2).$$

Это хорошо согласуется с интуицией: если состав шаров в черной и белой урнах одинаков, то вероятность выбрать черный шар не зависит от выбора урны.

Если, кроме того, в каждой из этих урн черных и белых шаров поровну ($b_1 = b_2 = c_1 = c_2$), то в примере 1 события A и B_1, A и B_2 независимы:

$$P_{B_1}(A) = P_{B_2}(A) = P(A) = 1/2.$$

Пример 3. *В условиях примера 2 пункта 2 из § 2 какова зависимость между цветом глаз отца и сына?*

Решение. В модели примера 2 пункта 2 из § 2 рассмотрим события A_1 и B_1, A_2 и B_2, описывающие соответственно темный и светлый цвет глаз у сыновей и отцов.

Из условий задачи следует, что

$$P(B_1) = 0{,}13; \quad P(B_2) = 0{,}87; \quad P_{B_1}(A_1) = 0{,}39;$$
$$P_{B_1}(A_2) = 0{,}61; \quad P_{B_2}(A_1) = 0{,}10; \quad P_{B_2}(A_2) = 0{,}90.$$

По правилу умножения и формуле полной вероятности получаем:

$$P(A_1B_1) = 0{,}13 \cdot 0{,}39 \approx 0{,}05;$$
$$P(A_1) = 0{,}13 \cdot 0{,}39 + 0{,}87 \cdot 0{,}10 \approx 0{,}14.$$

Следовательно,

$$K(A_1, B_1) \approx \frac{0{,}13 \cdot 0{,}39 - 0{,}13 \cdot 0{,}14}{\sqrt{0{,}14 \cdot 0{,}86 \cdot 0{,}13 \cdot 0{,}87}} \approx 0{,}30.$$

Так как события A_2 и B_2 дополнительны событиям A_1 и B_1, то

$$-K(A_1, B_2) = -K(A_2, B_1) = K(A_2, B_2) = K(A_1, B_1).$$

Таким образом, в рассматриваемой модели между цветом глаз отца и сына существует зависимость, оцениваемая коэффициентом корреляции с абсолютной величиной примерно 0,30. Для одинакового цвета этот коэффициент положителен, а для разного — отрицателен.

Пример 4. *В условиях примера 5 пункта 2 из § 2 какова зависимость между получением крысой пищи и ее поведением при повторном помещении в лабиринт?*

Решение. В модели примера 5 пункта 2 из § 2 рассмотрим события A и B, описывающие соответственно то, что крыса побывала в левом конце лабиринта при первом помещении ее туда и при втором. Из условий задачи следует, что

$$P(B) = 1/2; \quad P(B') = 1/2; \quad P_B(A) = 1/2 + \varepsilon; \quad P_{B'}(A) = 1/2.$$

По правилу умножения и формуле полной вероятности получаем:

$$P(AB) = 1/2(1/2+\varepsilon); \; P(A) = 1/2(1+\varepsilon).$$

Следовательно,

$$K(A, B) = \frac{(1/2)(1/2+\varepsilon) - (1/2)(1/2)(1+\varepsilon)}{\sqrt{(1/2)(1+\varepsilon)(1/2)(1-\varepsilon)(1/2)(1/2)}} =$$
$$= \frac{\varepsilon}{\sqrt{1-\varepsilon^2}}; \quad K(A', B) = -\frac{\varepsilon}{\sqrt{1-\varepsilon^2}}.$$

Если $\varepsilon = 0$, то $K(A, B) = 0$ и события A и B независимы. Если $\varepsilon \neq 0$, то $K(A, B) \neq 0$ и события A и B зависимы. В частности, если $\varepsilon = 1/2$, то $K(A, B) = 1/\sqrt{3}$.

Пример 5. *В условиях примера 7 пункта 2 из § 2 какова корреляция между посланным и принятым сигналами?*

Решение. В модели примера 7 пункта 2 из § 2 рассмотрим события A_1 и B_1, A_0 и B_0, описывающие соответственно принятый и посланный сигнал 1 и принятый и посланный сигнал 0. Из условий задачи следует, что

$$P(B_1) = p_1 = 0{,}6; \quad P(B_0) = p_0 = 0{,}4;$$

$$P_{B_1}(A_1) = r_{11} = 0{,}9; \quad P_{B_0}(A_1) = r_{01} = 0{,}3.$$

По правилу умножения и формуле полной вероятности получаем:

$$P(A_1B_1) = p_1r_{11} = 0{,}6\cdot 0{,}9 = 0{,}54;$$

$$P(A_1) = p_1r_{11} + p_{01}r_{01} = q_1 = 0{,}66.$$

Следовательно,

$$K(A_1, B_1) = \frac{p_1r_{11} - p_1q_1}{\sqrt{p_1p_0q_1q_0}} = \frac{0{,}6\cdot 0{,}9 - 0{,}6\cdot 0{,}66}{\sqrt{0{,}6\cdot 0{,}4\cdot 0{,}66\cdot 0{,}34}} \approx 0{,}6.$$

Так как события A_0 и B_0 дополнительны событиям A_1 и B_1, то

$$-K(A_0, B_1) = -K(A_1, B_0) = K(A_0, B_0) = K(A_1, B_1).$$

Замечание. Если $r_{11} = r_{01} = 1/2$, то $q_1 = q_0 = 1/2$, $K(A_1, B_1) = 0$ и события A_1 и B_1, A_0 и B_1, A_1 и B_0, A_0 и B_0 независимы. В этом случае можно сказать, что посылаемый и принимаемый сигналы независимы. Если $r_{11} = 1$ и $r_{01} = 0$, то $q_1 = p_1$ и $q_0 = p_0$, $K(A_1, B_1) = 1$, и в этом случае корреляция наибольшая.

Пример 6. *Проверяется эффективность нового медицинского препарата. Из имеющихся 60 зараженных животных*
30 вводится и 30 не вводится этот препарат.
Среди животных, которым был введен препарат,
29 выздоравливают и 1 нет.
Среди животных, которым не был введен препарат,
26 выздоравливают и 4 нет.
Какова корреляция между введением препарата и выздоровлением?

Решение. В модели, аналогичной рассматривавшимся в предыдущих примерах, дело сводится к вычислению коэффициента корреляции между событиями A и B, описывающими соответственно выздоровление и применение препарата. Из условий задачи следует, что

$$P(B) = 1/2; \quad P(B') = 1/2; \quad P_B(A) = 29/30; \quad P_{B'}(A) = 26/30.$$

По правилу умножения и формуле полной вероятности получаем:

$$P(AB) = (1/2)\,(29/30) = 29/60; \; P(A) = 55/60.$$

Следовательно,

$$K(A, B) = \frac{(1/2)(29/30) - (1/2)(55/60)}{\sqrt{(1/2)(1/2)(55/60)(5/60)}} = \frac{3}{5\sqrt{11}} \approx 0{,}2.$$

Пример 7. *В продукции завода брак вследствие дефекта A составляет 6,1%, а вследствие дефекта B — 2,8%. Общий брак по одному из этих дефектов — 5,8% всей продукции завода. Какова корреляция между дефектами A и B?*

Решение. В модели, аналогичной рассматривавшимся в предыдущих примерах, дело сводится к вычислению коэффициента корреляции между событиями A и B, описывающими соответствующие дефекты. Из условий задачи следует, что

$$P(A) = 0{,}061; \quad P(B) = 0{,}028; \quad P(A \cup B) = 0{,}058.$$

Используя правило объединения, находим:

$$P(AB) = P(A) + P(B) - P(A \cup B) = 0{,}030.$$

Следовательно,

$$K(A, B) = \frac{0{,}030 - 0{,}061 \cdot 0{,}028}{\sqrt{0{,}061 \cdot 0{,}939 \cdot 0{,}028 \cdot 0{,}972}} \approx 0{,}832.$$

Пример 8. *Статистика показывает, что среди двоен 32% оба близнеца мальчики и 28% — девочки. Какова корреляция пола близнецов?*

Решение. В модели, аналогичной рассматривавшимся в предыдущих примерах, дело сводится к вычислению коэффициента корреляции событий A и B, описывающих мужской пол одного и другого близнецов. Из условий задачи следует, что

$$P(AB) = 0{,}32; \; P(A'B') = 0{,}28.$$

Используя правила сложения и дополнения, находим:

$$P(A'B + AB') = P(A'B) + P(AB') = P(B) - P(AB) + P(B') - \\ - P(A'B') = 1 - 0{,}32 - 0{,}28 = 0{,}40.$$

Естественно считать, что

$$P(A'B) = P(AB') = 0{,}20.$$

Тогда

$$P(A) = P(AB) + P(AB') = 0{,}52; \; P(B) = P(AB) + P(A'B) = 0{,}52.$$

Следовательно,

$$K(A, B) = \frac{0{,}32 - 0{,}52 \cdot 0{,}52}{\sqrt{0{,}52 \cdot 0{,}48 \cdot 0{,}52 \cdot 0{,}48}} \approx 0{,}2.$$

Кроме того,

$$-K(A', B) = -K(A, B') = K(A', B') = K(A, B).$$

§ 4. РАЗНЫЕ ЗАДАЧИ

Рассмотрим несколько задач различного содержания.

1.4. Задача де Мере

Сколько раз нужно подбрасывать 2 игральные кости, чтобы вероятность хотя бы один раз получить 2 шестерки была больше половины?

Решение. Предположим, что кости разноцветные: красная и белая.

Результаты одного подбрасывания такой пары костей можно описать множеством всех 36 пар u_1u_2, составленных из чисел 1, 2, 3, 4, 5, 6. Первое число u_1 описывает число очков на красной кости, а второе число u_2 — на белой. Результаты m последовательных подбрасываний можно описать множеством U всех строк

$$U=(u_{11}u_{21},\ u_{12}u_{22},\ \ldots,\ u_{1m}u_{2m})$$

длины m, составленных из пар чисел 1, 2, 3, 4, 5, 6. Первое число i-й пары u_{1i} описывает число очков на красной кости при i-м подбрасывании, а второе число u_{2i} — на белой ($i=1, \ldots, m$). По правилу умножения таких строк 36^m.

Симметричность игральных костей позволяет считать все результаты равновозможными. Поэтому для описания рассматриваемого опыта можно использовать модель Лапласа 36^m равновероятных исходов.

Задача сводится к вычислению вероятности $P(A)$ события A, составленного из всех строк u, в которых есть хотя бы одна пара 66. Это событие описывает появление хотя бы при одном подбрасывании двойной шестерки: 6 очков на красной кости и 6 — на белой.

Проще вычислить сначала вероятность дополнительного события A', составленного из всех строк u, в которых нет ни одной пары 66. По правилу умножения таких строк 35^m.
Следовательно,

$$P(A')=n(A')/n(U)=35^m/36^m=(35/36)^m.$$

По правилу дополнения отсюда вытекает, что

$$P(A)=1-P(A')=1-(35/36)^m,$$

неравенство

$$P(A)\geqslant 1/2$$

поэтому эквивалентно неравенству

$$(35/36)^m\leqslant 1/2.$$

В свою очередь это неравенство эквивалентно неравенству

$$m\geqslant \log(1/2)/\log(35/36)\approx 24{,}6.$$

Таким образом, для того, чтобы вероятность появления двойной шестерки была больше половины, нужно подбрасывать кости самое меньшее 25 раз.

Замечание. Задача де Мере является одной из первых задач, с которыми связано зарождение современной теории вероятностей. В середине XVII века любивший азартные игры французский дворянин де Мере предложил эту задачу одному из выдающихся ученых того времени Паскалю.

Задача де Мере возникла в связи со следующей игрой. Две кости подбрасываются 24 раза. Можно ставить либо на появление хотя бы один раз *двойной шестерки*, либо против этого результата. Проведенные рассуждения показывают, что в такой игре на двойную шестерку ставить невыгодно: вероятность выигрыша в этом случае равна

$$1-(35/36)^{24}=0{,}491404<1/2.$$

Сначала де Мере ставил на появление хотя бы одной шестерки при подбрасывании одной кости 4 раза и, как правило, выигрывал чаще, чем проигрывал. (Вероятность появления хотя бы одной шестерки при четырех подбрасываниях одной кости равна $1-(5/6)^4=$ $=671/1296>1/2$). Когда это было замечено, де Мере начал ставить на появление хотя бы одной пары шестерок при подбрасывании двух костей 24 раза и, как правило, чаще проигрывал, чем выигрывал ($1-(35/36)^{24}<1/2$).

Сам де Мере правильно подсчитал, что вероятность появления двойной шестерки при подбрасывании пары костей в 6 раз меньше вероятности появления шестерки при подбрасывании одной кости. Отсюда он сделал неправильный вывод о том, что вероятность q появления двойной шестерки при четырех подбрасываниях пары костей в 6 раз меньше вероятности p появления шестерки при четырех подбрасываниях одной кости: $q=1/6\cdot p$, а вероятность появления двойной шестерки при $6\cdot 4=24$ подбрасываниях пары костей равна $6\cdot q=p>1/2$.

2.4. Задача о красных шарах

Имеются n шаров, l из которых красные, а $n-l$ — белые. Из этих n шаров наугад выбираются m. Какова вероятность того, что среди выбранных m шаров ровно k оказываются красными?

Решение. Предположим, что шары отмечены номерами $1, \ldots, n$, причем красные имеют номера $1, \ldots, l$, а белые — $l+1, \ldots, n$ $(1\leqslant l<n)$. Результаты выбора m из этих n шаров можно описать множеством U всех выборок

$$u=\{u_1, \ldots, u_m\} \qquad (1\leqslant m\leqslant n)$$

m номеров из n номеров $1, \ldots, n (1 \leqslant m \leqslant n)$. Выборка u описывает выбор шаров с номерами $u_1, \ldots, u_m$. Число всех таких выборок равно $\binom{n}{m}$.

Условие о выборе шаров *наугад* позволяет считать все результаты равновозможными. Поэтому для описания рассматриваемого опыта можно использовать модель Лапласа $\binom{n}{m}$ равновероятных исходов.

Задача сводится к вычислению вероятности $P(A)$ события A, составленного из всех выборок, содержащих ровно k из номеров $1, \ldots, l$ $(0 \leqslant k \leqslant l)$. Существуют ровно $\binom{l}{k}$ выборок k из номеров $1, \ldots, l$ и ровно $\binom{n-l}{m-k}$ выборок $(m-k)$ из номеров $l+1, \ldots, n$. По правилу умножения отсюда вытекает, что существует ровно

$$n(A) = \binom{l}{k}\binom{n-l}{m-k}$$

выборок m из номеров $1, \ldots, n$, содержащих ровно k из номеров $1, \ldots, l$. Следовательно,

$$P(A) = \binom{l}{k}\binom{n-l}{m-k} \Big/ \binom{n}{m}.$$

Задача о красных шарах имеет многочисленные приложения.

1.2.4. Статистический контроль

Имеются $n=100$ изделий, $l=2$ из которых негодные, а $n-l=98$ — годные. Из этих $n=100$ изделий наугад выбираются $m=10$. Какова вероятность того, что среди выбранных $m=10$ изделий ровно $k=1$ оказываются негодными?

Дело сводится к задаче о красных шарах, которые играют роль негодных изделий. Искомая вероятность равна

$$\binom{2}{1}\binom{n-2}{m-1} \Big/ \binom{n}{m} = 2\frac{m}{n}\cdot\frac{n-m}{n-1} = 2/11.$$

2.2.4. Тип гена

Предположим, что *ген* состоит из $n_0=3$ частей, из которых $l_0=1$ *мутантные*, а $n_0-l_0=2$ — *немутантные*. Перед делением клетки части удваиваются, причем каждая мутантная дает *две мутантные* и каждая немутантная — *две немутантные*. При делении образуются две новые клетки, содержащие по одному новому гену, составленному $m=n_0$ частями из имевшихся $n=2n_0$. Про-

цесс образования новых генов описывается выбором наугад m из имеющихся n частей. *Тип гена* определяется числом k мутантных частей в нем.

Задача. *Какова вероятность того, что при делении клетки с геном типа $l_0=1$, состоящим из $n_0=3$ частей, получится клетка с геном типа $k=2$?*

Решение. Дело сводится к задаче о красных шарах при $n=2n_0=6$, $l=2l_0=2$, $m=n_0=3$, $k=2$. Красные шары играют роль мутантных частей гена. Искомая вероятность равна

$$\binom{2}{2}\binom{6-2}{3-2}\Big/\binom{6}{3}=1/5.$$

3.2.4. Анализ крови

В данном объеме крови имеются $n=5\cdot 10^6+5\cdot 10^3$ кровяных телец, $l=5\cdot 10^6$ которых красные, а $n-l=5\cdot 10^3$ — белые. Из этих n телец наугад выбираются $m=10$. Какова вероятность того, что среди выбранных $m=10$ кровяных телец ровно $k=0$ оказываются красными?

Решение. Дело сводится к задаче о красных шарах, которые играют роль красных кровяных телец. Искомая вероятность $P(A)$ равна

$$P(A)=\binom{l}{0}\binom{n-l}{m-0}\Big/\binom{n}{m}=\binom{n-l}{m}\Big/\binom{n}{m}=$$
$$=\frac{(n-l)(n-l-1)\ldots(n-l-m+1)}{n(n-1)\ldots(n-m+1)}.$$

Нетрудно убедиться в том, что эта вероятность чрезвычайно мала. Действительно, каждый множитель в числителе меньше числа $n-l=5\cdot 10^3$, а в знаменателе — больше числа $n-m+1=5\cdot 10^6+5\cdot 10^3-9>5\cdot 10^6$. Следовательно,

$$P(A)\leqslant\left(\frac{5\cdot 10^3}{5\cdot 10^6}\right)^{10}=10^{-30}.$$

Этот результат интуитивно ясен: так как красные кровяные тельца составляют подавляющее большинство, то чрезвычайно мало вероятно, что их не окажется в случайной выборке.

4.2.4. Метод меченых частиц

Имеются n частиц, l из которых отмечены, а $n-l$ — не отмечены. Из этих n частиц наугад выбираются m. Среди них оказывается ровно k отмеченных. При каком числе n частиц это наиболее вероятно?

Решение. Дело сводится к задаче о красных шарах, играющих роль отмеченных частиц. Нужно найти $n=\hat{n}$, при котором вероятность

$$P(A)=\binom{l}{k}\binom{n-l}{m-k}\Big/\binom{n}{m}$$

наибольшая. Для каждого $n>m+l-k$ отношение

$$\frac{\binom{l}{k}\binom{n-l}{m-k}\Big/\binom{n}{m}}{\binom{l}{k}\binom{n-1-l}{m-k}\Big/\binom{n-1}{m}}=\frac{(n-l)(n-m)}{n(n-m-l+k)}$$

больше 1, если $nk<lm$, и меньше 1, если $nk>lm$. Следовательно, если $k>0$, то целая часть

$$\hat{n}=[ml/k]$$

числа ml/k является *наиболее правдоподобной оценкой* для числа частиц: при $n=\hat{n}$ и данных m, l, k вероятность $P(A)$ наибольшая.

Задача о рыбах. *Из озера вылавливается $l=1000$ рыб. Каждая из них метится и выпускается в озеро. Затем снова вылавливается $m=1000$ рыб. Среди них оказывается $k=100$ меченых. Каково наиболее правдоподобное число рыб в озере?*

Решение. Предположим, что второй улов организован так, что его можно считать выбором наугад m из имеющихся в озере n рыб. В этом случае наиболее правдоподобной оценкой для числа n рыб в озере является число

$$\hat{n}=1000\cdot 1000/100=10\,000.$$

5.2.4. Статистический контроль

Имеются n изделий, l из которых негодные, а $n-l$ — годные. Из этих n изделий наугад выбираются m. Среди них оказывается ровно k негодных. При каком числе l негодных изделий это наиболее вероятно?

Решение. Дело сводится к задаче о красных шарах, играющих роль негодных изделий. Нужно найти число $l=\hat{l}$, при котором вероятность

$$P(A)=\binom{l}{k}\binom{n-l}{m-k}\Big/\binom{n}{m}$$

наибольшая. Для каждого $l\neq k$ отношение

$$\frac{\binom{l}{k}\binom{n-l}{m-k}\Big/\binom{n}{m}}{\binom{l-1}{k}\binom{n-l+1}{m-k}\Big/\binom{n}{m}}=\frac{l(n-m-l+k+1)}{(l-k)(n-l+1)}$$

больше 1, если $lm<(n+1)k$, и меньше 1, если $lm>(n+1)k$. Следовательно, целая часть

$$\hat{l}=[(n+1)k/m]$$

числа $(n+1)k/m$ является *наиболее правдоподобной оценкой* для числа l негодных изделий: при $l=\hat{l}$ и данных n, m, k вероятность $P(A)$ наибольшая.

Задача. *Из партии $n=999$ гвоздей наугад выбирается $m=100$ гвоздей. Среди них оказывается $k=10$ негодных. Каково наиболее правдоподобное число негодных гвоздей в партии?*

Ответ: наиболее правдоподобной оценкой для числа l негодных гвоздей в партии является число

$$\hat{l}=1000\cdot 10/100=100.$$

3.4. Задача о размещении

Шары случайным образом размещаются по ящикам. Какова вероятность того, что в каждом ящике оказывается данное число шаров?

Ответ зависит от того, каким образом размещаются шары по ящикам.

1.3.4. Различные шары

Рассмотрим m различных шаров $1, \ldots, m$ и n различных ящиков $1, \ldots, n$. Предположим, что каждый шар кладется в наугад выбранный ящик. В этом случае задачу о размещении можно сформулировать следующим образом.

Задача 1. *Каждый из шаров $1, \ldots, m$ кладется в наугад выбранный из ящиков $1, \ldots, n$. Какова вероятность того, что в первом ящике оказывается m_1 шаров, ..., а в n-м — m_n шаров?*

Решение. Рассмотрим произвольные натуральные числа $m\geqslant 1$, $n\geqslant 1$ и $m_1, \ldots, m_n$:

$$m_1+\ldots+m_n=m.$$

Каждое размещение шаров $1, \ldots, m$ по ящикам $1, \ldots, n$ можно описать строкой $u=u_1\ldots u_m$ длины m, составленной из номеров $1, \ldots, n$: шар i кладется в ящик u_i. Из правила умножения для числа элементов следует, что таких строк n^m.

Условие о выборе наугад позволяет считать все размещения шаров равновозможными и использовать модель Лапласа n^m равновероятных исходов.

Задача сводится к вычислению вероятности $P(A)$ события A, составленного из всех строк, в которых m_1 номеров $1, \ldots, m_n$ номеров n.

Используя принцип индукции, нетрудно проверить, что

$$n(A)=\frac{m!}{m_1!\dots m_n!}.$$

Действительно, если $n=1$, то это равенство верно: в этом случае множество A состоит из единственной строки, составленной из единиц. Для каждого номера n, если равенство верно для строк номеров $1, \dots, n$, то оно верно для строк номеров $1, \dots, n, n+1$. В самом деле, множество таких строк можно разбить на части, состоящие из строк с одинаково расположенными m_{n+1} номерами $n+1$. Таких частей $\binom{m}{m_{n+1}}$. Если число строк в каждой части выбирается рассматриваемым равенством для $m-m_{n+1}$, то, по правилу сложения, число рассматриваемых строк номеров $1, \dots, n, n+1$ равно

$$n(A)=\frac{(m-m_{n+1})!}{m_1!\dots, m_n!}\binom{m}{m_{n+1}}=$$

$$=\frac{(m-m_{n+1})!}{m_1!\dots m_n!}\cdot\frac{m!}{(m_{n+1})!\,(m-m_{n+1})!}=\frac{m!}{m_1!\dots m_{n+1}!}.$$

Таким образом,

(1) $$P(A)=\frac{m!}{m_1!\dots m_n!}n^{-m}.$$

2.3.4. Одинаковые шары

Рассмотрим m одинаковых шаров и n различных ящиков $1, \dots, n$. Предположим, что размещения шаров по ящикам равновозможны. В этом случае задачу о размещении можно сформулировать следующим образом.

Задача 2. *Наугад выбирается размещение m одинаковых шаров по n различным ящикам $1, \dots, n$. Какая вероятность того, что в ящике 1 оказывается m_1 шаров, ..., а в ящике $n-m_n$ шаров?*

Решение. Рассмотрим произвольные натуральные числа $m\geqslant 1$, $n\geqslant 1$ и $m_1, \dots, m_n$:

$$m_1+\dots+m_n=m.$$

Размещение m одинаковых шаров по различным ящикам $1, \dots, n$, при котором в ящике 1 оказывается m_1 шаров, ..., а в ящике $n-m_n$ шаров, можно описать строкой u длины $m+n$, составленной из m одинаковых кружочков и номеров $1, \dots, n$.

Первым элементом строки u является номер 1, следующие m_1 элементов — кружочки, ..., $n+(m_1+m_2+\dots+m_{n-1})$-м элементом — номер n, следующие m_n элементов — кружки (рис. 37).

$$u=1\,\underbrace{0\dots0}_{m_1}\,2\,\underbrace{0\dots0}_{m_2}\,\dots\,n\,\underbrace{0\dots0}_{m_n}$$

Рис. 37.

Выбор размещений наугад позволяет считать все размещения шаров равновозможными и использовать модель Лапласа.

Задача сводится к вычислению числа всех рассматриваемых строк u. Каждая такая строка определяется местами кружочков в ней. Поэтому число всех строк равно числу выборок m из номеров 2, ..., $m+n$. Таким образом,

(2)
$$p(u) = 1 \Big/ \binom{n+m-1}{m}.$$

Замечание. Искомая вероятность выражается также равенством

$$p(u) = 1 \Big/ \binom{n+m-1}{n-1},$$

соответствующим выбору мест для номеров 2, ..., n в строке u.

Задача о размещении шаров имеет многочисленные приложения. Шары и ящики могут, например, играть роли соответственно:

1) числа очков и номеров подбрасывания игральной кости;
2) карт и их мастей;
3) частиц и областей пространства;
4) людей и их дней рождения;
5) биологических особей и их генотипов;
6) молекул и цепочек молекул.

Пример 1. *Игральная кость подбрасывается шесть раз. Какова вероятность того, что каждый раз появляется новое число очков?*

Решение. Дело сводится к задаче о размещении для различных шаров, играющих роль числа очков. Ящики играют роль подбрасываний кости. Используя равенство (1), для $m=n=6$ и $m_1= \ldots =m_6=1$ находим, что искомая вероятность равна

$$6! \cdot 6^{-6} = 0{,}01543.$$

Пример 2. *Тщательно перемешанная колода из 52 игральных карт делится поровну между четырьмя игроками. Какова вероятность того, что у данного игрока оказываются карты каждой масти?*

Решение. Дело сводится к задаче о размещении $m=13$ одинаковых шаров, играющих роль карт. Ящики 1, 2, 3, 4 играют роль мастей. Нужно вычислить вероятность того, что ни один из ящиков не оказывается пустым. Отсутствие пустых ящиков описывает событие A, составленное из всех строк u, в которых каждый из $n-1=3$ номеров 2, 3, 4 занимает одно из $m-1=12$ мест между кружочками. Например, $u=10\,200\,300\,040\,000\,000$.

Число таких строк равно $\binom{12}{3}$. Следовательно,

$$P(A) = \binom{12}{3} \Big/ \binom{4+13-1}{13} = 11/28.$$

Пример 3. Модель Максвелла — Больцмана. Рассмотрим m *различных частиц*, распределенных по n *областям пространства.* Предположим, что все распределения частиц *равновозможны*, а состояние системы определяется указанием числа частиц, попавших в каждую область пространства.

Условимся описанную модель называть *моделью Максвелла — Больцмана.* Модель Максвелла — Больцмана сводится к модели *различных шаров*, если частицы считать шарами, а области пространства — ящиками. Поэтому в модели Максвелла — Больцмана вероятность того, что система оказывается в состоянии $(m_1, \ldots, m_n)$, выражается равенством (1).

Задача. *Какова в модели Максвелла — Больцмана вероятность того, что в каждой области пространства оказывается ровно одна частица?*

Ответ: $n!/n^n$. Когда n увеличивается, эта вероятность уменьшается. Если $n \geqslant 10$, то $n!/n^n \leqslant 0{,}0004$.

Пример 4. Модель Бозе — Эйнштейна. Рассмотрим m *одинаковых частиц*, распределенных по n *областям пространства.* Предположим, что все распределения равновозможны, а состояние системы определяется указанием числа частиц, попавших в каждую область пространства.

Условимся описанную модель называть *моделью Бозе — Эйнштейна.* Модель Бозе — Эйнштейна сводится к модели *одинаковых шаров*, если частицы считать шарами, а области пространства — ящиками. Поэтому в модели Бозе — Эйнштейна вероятность того, что система оказывается в состоянии $(m_1, \ldots, m_n)$, выражается равенством (2).

Задача. *Какова в модели Бозе — Эйнштейна вероятность того, что в каждой области пространства оказывается хотя бы одна частица?*

Решение. В модели m одинаковых шаров, распределенных по n ящикам, дело сводится к вычислению вероятности того, что не оказывается пустых ящиков. Отсутствие пустых ящиков описывает событие A, составленное из всех строк u, в которых нет рядом стоящих номеров: каждый из $n-1$ номеров $2, \ldots, n$ занимает одно из $m-1$ мест между кружочками. Поэтому

$$n(A) = \binom{m-1}{n-1}, \qquad P(A) = \binom{m-1}{n-1} \Big/ \binom{n+m-1}{n-1}.$$

В частности, если $m < n$, то эта вероятность равна нулю: ящиков больше, чем шаров, и обязательно останутся пустые.

4.4. Задача о крэпсе

Подбрасываются две игральные кости. Если сумма очков равна 7 или 11, то игрок выигрывает; если 2 или 3 или 12,— проигрывает. При каждой другой сумме первая партия окан-

чивается вничью и кости подбрасываются снова и снова до повторения первой суммы или до суммы 7 (но не больше, чем n раз). В первом случае игрок выигрывает, а во втором — проигрывает, В других случаях игра оканчивается вничью. Какова вероятность выигрыша игрока?

Решение. Построим вероятностную модель для рассматриваемой игры. В качестве множества исходов возьмем множество U всех строк $u=u_1\ldots u_j\ldots u_n$ длины n, составленных из чисел 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12. Число s на j-м месте строки $u(u_j=s)$ описывает сумму s при j-м подбрасывании. Используя зависимость подбрасываний, определим элементарную вероятность p следующим образом, При решении задачи 4 пункта 1 из § 1 было показано, что одно подбрасывание двух костей описывается элементарной вероятностью q, значения $q(s)$ которой содержит приведенная там таблица. Значение $p(u)$ для каждой строки $u=u_1\ldots$ $\ldots u_j\ldots u_n$ элементарной вероятности p, описывающей n подбрасываний двух костей, определим равенством

$$p(u)=q(u_1)\ldots q(u_j)\ldots q(u_n).$$

Для каждого номера $j=1, \ldots, n$ рассмотрим события A^j, B^j и C^j, описывающие соответственно *выигрыш при j-й партии, проигрыш при j-й партии* и *ничью при каждой из первых j партий.*

В частности,

$A^1=\{u:u_1=7,11\}$, $B^1=\{u:u_1=2, 3, 12\}$, $C^1=\{u:u_1=4, 5, 6, 8, 9, 10\}$, $A^2=\{u:u_1=u_2=4\}+\{u:u_1=u_2=5\}+\{u:u_1=u_2=6\}+\{u:u_1=u_2=8\}+\{u:u_1=u_2=9\}+\{u:u_1=u_2=10\}$, $B^2=\{u:u_1=4, u_2=7\}+\{u:u_1=5, u_2=7\}+\{u:u_1=6, u_2=7\}+\{u:u_1=8, u_2=7\}+$

$$+\{u:u_1=8, u_2=7\}+\{u:u_1=9, u_2=7\}+\{u:u_1=10, u_2=7\},$$

$$C^2=\{u:u_1=4, u_2\neq 4,7\}+\{u:u_1=5, u_2\neq 5,7\}+\{u:u_1=6, u_2\neq 6,7\}+$$
$$+\{u:u_1=8, u_2\neq 8,7\}+\{u:u_1=9, u_2\neq 9,7\}+\{u:u_1=10, u_2\neq 10,7\}.$$

Если $3\leqslant j\leqslant n$, то

$A^j=\{u:u_1=4, u_2\neq 4,7, \ldots, u_{j-1}\neq 4,7, u_j=4\}+\{u:u_1=5, u_2\neq 5,7, \ldots, u_{j-1}\neq 5,7, u_j=5\}+\{u:u_1=6, u_2\neq 6,7, \ldots, u_{j-1}\neq 6,7, u_j=6\}+$ $+\{u:u_1=8, u_2\neq 8,7, \ldots, u_{j-1}\neq 8,7, u=8\}+\{u:u_1=9, u_2\neq 9,7, \ldots$ $\ldots, u_{j-1}\neq 9,7, u_j=9\}+\{u:u_1=10, u_2\neq 10,7, \ldots, u_{j-1}\neq 10,7, u_j=10\}$,

$B^j=\{u:u_1=4, u_2\neq 4,7, \ldots, u_{j-1}\neq 4,7, u_j=7\}+\{u:u_1=5, u_2\neq 5,7, \ldots, u_{j-1}\neq 5,7, u_j=7\}+\{u:u_1=6, u_2\neq 6,7, \ldots, u_{j-1}\neq 6,7, u_j=7\}+$ $+\{u:u_1=8, u_2\neq 8,7, \ldots, u_{j-1}\neq 8,7, u_j=7\}+\{u:u_1=9, u_2\neq 9,7, \ldots,$ $u_{j-1}\neq 9,7, u_j=7\}+\{u:u_1=10, u_2\neq 10,7, \ldots, u_{j-1}\neq 10,7, u_j=7\}$.

Ясно, что события $A^1, \ldots, A^n, B^1, \ldots, B^n$ попарно не пересекаются. Их суммы

$$A=A^1+\ldots+A^n,\ B=B^1+\ldots+B^n$$

описывают соответственно выигрыш и проигрыш игрока, когда игра продолжается не более, чем n партий. Пересечение $C=C^1\ldots$ $\ldots C^n$ событий $C^1, \ldots, C^n$ описывает ничью. Ясно, что события A, B и C попарно не пересекаются и

$$A+B+C=U.$$

Из *правила сложения* следует, что

$$P(A)=P(A^1)+\ldots+P(A^n),\ P(B)=P(B^1)+\ldots P(B^n),$$

$$P(C)=1-P(A)-P(B).$$

Таким образом, дело сводится к вычислению вероятностей событий A^j и B^j.

Используя определение элементарной вероятности p и соответствующую таблицу, находим:

$$P(A^1)=q(7)+q(11)=2/9;\ P(B^1)=q(2)+q(3)+q(12)=1/9;$$

$$P(A^2)=q^2(5)+q^2(4)+q^2(6)+q^2(8)+q^2(10)=100/(36)^2;$$

$$P(B^2)=[q(4)+q(5)+q(6)+q(8)+q(9)+q(10)]q(7)=144/(36)^2.$$

Заметим, что

$$q(4)=q(10)=3/36;\ q(9)=q(5)=4/36;\ q(6)=q(8)=5/36;$$

$$q(2)+q(3)+q(5)+q(6)+q(8)+q(9)+q(10)+q(11)+$$
$$+q(12)=27/36;$$

$$q(2)+q(3)+q(4)+q(6)+q(8)+q(9)+q(10)+q(11)+$$
$$+q(12)=26/36;$$

$$q(2)+q(3)+q(5)+q(8)+q(9)+q(10)+q(11)+q(12)=25/36.$$

Если $3\leqslant j\leqslant n$, то, используя правило сложения, определение элементарной вероятности p и эти равенства, получаем:

$$P\left(A^j\right)=2\left[\frac{3}{36}\left(\frac{27}{36}\right)^{j-2}\frac{3}{36}+\frac{4}{36}\left(\frac{26}{36}\right)^{j-2}\frac{4}{36}+\frac{5}{36}\left(\frac{25}{36}\right)^{j-2}\frac{5}{36}\right];$$

$$P\left(B^j\right)=2\left[\frac{3}{36}\left(\frac{27}{36}\right)^{j-2}\frac{6}{36}+\frac{4}{36}\left(\frac{26}{36}\right)^{j-2}\frac{6}{36}+\frac{5}{36}\left(\frac{25}{36}\right)^{j-2}\frac{6}{36}\right].$$

Используя правило сложения и эти равенства, находим:

$$P(A)=\frac{2}{9}+2\left[\left(\frac{3}{36}\right)^2\frac{1-(27/36)^{n-1}}{1-(27/36)}+\left(\frac{4}{36}\right)^2\frac{1-(26/36)^{n-1}}{1-(26/36)}+\right.$$
$$\left.+\left(\frac{5}{36}\right)^2\frac{1-(25/36)^{n-1}}{1-(25/36)}\right]=\frac{244}{295}-r(A);$$

$$P(B)=\frac{1}{9}+\frac{2\cdot 6}{36}\left[\frac{3}{36}\cdot\frac{1-(27/36)^{n-1}}{1-(27/36)}+\frac{4}{36}\cdot\frac{1-(26/36)^{n-1}}{1-(26/36)}+\right.$$
$$\left.+\frac{5}{36}\cdot\frac{(1-25/36)^{n-1}}{1-(25/36)}\right]=\frac{251}{495}-r(B),$$

где

$$r(A) = \frac{2}{36}\left[\left(\frac{27}{36}\right)^{n-1} + \frac{8}{5}\left(\frac{26}{36}\right)^{n-1} + \frac{25}{11}\left(\frac{25}{36}\right)^{n-1}\right];$$

$$r(B) = \frac{1}{3}\left[\frac{1}{3}\left(\frac{27}{36}\right)^{n-1} + \frac{2}{5}\left(\frac{26}{36}\right)^{n-1} + \frac{5}{11}\left(\frac{25}{36}\right)^{n-1}\right].$$

Отсюда следует, что

$$P(C) = r(A) + r(B) = \frac{1}{18}\left[3\left(\frac{27}{36}\right)^{n-1} + 4\left(\frac{26}{36}\right)^{n-1} + 5\left(\frac{25}{36}\right)^{n-1}\right].$$

Заметим, что

$$r(A) \leqslant \frac{1}{18}\left[1 + \frac{8}{5} + \frac{25}{11}\right]\left(\frac{27}{36}\right)^{n-1} \leqslant \frac{1}{3}\left(\frac{3}{4}\right)^{n-1};$$

$$r(B) \leqslant \frac{1}{3}\left[\frac{1}{3} + \frac{2}{5} + \frac{5}{11}\right]\left(\frac{27}{36}\right)^{n-1} \leqslant \frac{1}{2}\left(\frac{3}{4}\right)^{n-1};$$

$$P(C) \leqslant \frac{1}{18}[3 + 4 + 5]\left(\frac{27}{36}\right)^{n-1} \leqslant \frac{2}{3}\left(\frac{3}{4}\right)^{n-1}.$$

При достаточно большом числе n партий эти величины произвольно малы. Если n больше 25, то каждая из них меньше 0,0018.

Таким образом, при достаточно большом числе n партий ($n \geqslant 25$) вероятности выигрыша, проигрыша и ничьей выражаются соответственно приближенными равенствами

$$P(A) \approx \frac{244}{495} \approx 0{,}493; \quad P(B) \approx \frac{251}{495} \approx 0{,}507; \quad P(C) \approx 0.$$

Вероятность выигрыша приблизительно на 0,014 меньше вероятности проигрыша.

Замечание. Рассматривавшаяся игра называется крэпс и очень популярна в некоторых странах.

5.4. Задача о красных, белых и розовых урнах

В каждой из комнат 1 и 2 имеются:

1) красные урны 1, ..., a, каждая с красным шаром 1 и красным шаром 2;

2) розовые урны a+1, ..., a+b, каждая с красным шаром 1 и белым шаром 2;

3) розовые урны a+b+1, ..., a+2b, каждая с белым шаром 1 и красным шаром 2;

4) белые урны a+2b+1, ..., a+2b+c, каждая с белым шаром 1 и белым шаром 2;

В каждой из комнат наугад выбирается урна. В каждой из этих урн наугад выбирается шар. Какова вероятность того, что оба выбираемые шара белые?

Решение. Возможные результаты рассматриваемого опыта удобно описывать матрицами

$$u = \begin{pmatrix} u_{10} & u_{11} & u_{12} & u_{13} \\ u_{20} & u_{21} & u_{22} & u_{23} \end{pmatrix},$$

в которых $(i=1,\ 2)$: u_{i0} равно *номеру* урны, выбранной в комнате i; $u_{i1}=1$, если в ней шар 1 *красный*, и $=0$, если *белый*; $u_{i2}=1$, если в ней шар 2 *красный*, и $=0$, если *белый*; $u_{i3}=1$, если выбранный из нее шар *красный*, и $=0$, если *белый*.

Эти матрицы u образуют множество исходов U.

Условие о выборе наугад урн и шаров позволяет использовать классическую модель Лапласа.

Рассмотрим события

$$B\begin{pmatrix} ij \\ lm \end{pmatrix} = \left\{u : \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} = \begin{pmatrix} ij \\ lm \end{pmatrix}\right\} \quad (i, j, l, m = 0, 1),$$

составленные из матриц u со средними столбцами $\begin{pmatrix} ij \\ lm \end{pmatrix}$ и описывающие выбираемые *урны*, а также события

$$A\begin{pmatrix} k \\ n \end{pmatrix} = \left\{u : \begin{pmatrix} u_{13} \\ u_{23} \end{pmatrix} = \begin{pmatrix} k \\ n \end{pmatrix}\right\} \quad (k, n = 0, 1),$$

составленные из матриц u с последним столбцом $\begin{pmatrix} k \\ n \end{pmatrix}$ и описывающие выбираемые шары. Задача сводится к вычислению вероятности события $A\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

По формуле полной вероятности

$$P\left[A\begin{pmatrix} k \\ n \end{pmatrix}\right] = \sum_{i,j,l,m} P\left[\left(B\begin{pmatrix} ij \\ lm \end{pmatrix}\right] P_{B\binom{ij}{lm}}\left[A\begin{pmatrix} k \\ n \end{pmatrix}\right].$$

Таким образом, дело сводится к вычислению вероятностей событий $B\begin{pmatrix} ij \\ lm \end{pmatrix}$ и условных вероятностей событий $A\begin{pmatrix} k \\ n \end{pmatrix}$.

Положим:

$$\bar{a} = \frac{a}{a+2b+c}, \quad \bar{b} = \frac{b}{a+2b+c}, \quad \bar{c} = \frac{c}{a+2b+c}.$$

Как нетрудно проверить, вероятности событий $B\begin{pmatrix} ij \\ lm \end{pmatrix}$ содержит таблица

$ij \setminus lm$	11	10	01	00
11	$\bar{a}\bar{a}$	$\bar{a}\bar{b}$	$\bar{a}\bar{b}$	$\bar{a}\bar{c}$
10	$\bar{b}\bar{a}$	$\bar{b}\bar{b}$	$\bar{b}\bar{b}$	$\bar{b}\bar{c}$
01	$\bar{b}\bar{a}$	$\bar{b}\bar{b}$	$\bar{b}\bar{b}$	$\bar{b}\bar{c}$
00	$\bar{c}\bar{a}$	$\bar{c}\bar{b}$	$\bar{c}\bar{b}$	$\bar{c}\bar{c}$

Условные вероятности событий $A\binom{k}{m}$ содержит таблица

k \ ij n \ lm	11 11	11 10	11 01	11 00	10 11	10 10	10 01	10 00	01 11	01 10	01 01	01 00	00 11	00 10	00 01	00 00
1 1	1	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	0	0	0	0	0
1	0	$\frac{1}{2}$	$\frac{1}{2}$	1	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	0	0	0	0
0 1	0	0	0	0	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	0	1	$\frac{1}{2}$	$\frac{1}{2}$	0
0 0	0	0	0	0	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$	1

Используя таблицы и ***формулу полной вероятности***, находим:

$$P\left[A\binom{1}{1}\right] = (\bar{a} + \bar{b})^2;$$

$$P\left[A\binom{1}{0}\right] = P\left[A\binom{0}{1}\right] = (\bar{a} + \bar{b})(\bar{b} + \bar{c});$$

$$P\left[A\binom{0}{0}\right] = (\bar{b} + \bar{c})^2.$$

Таким образом, ответ:

$$P\left[A\binom{0}{0}\right] = \left(\frac{b + c}{a + 2b + c}\right)^2.$$

Замечание. Задача о красных, белых и розовых урнах имеет отношение к генетической модели наследования некоторого признака. По этому признаку каждая *особь* рассматриваемой популяции относится к одному из трех возможных *генотипов* gg, GG и $\{gG, Gg\}$. В первых двух случаях говорят о *чистом* генотипе, а в последнем — о *смешанном*.

Предполагается, что каждая мужская и женская особи, образуя пару, порождают одного *потомка*. В соответствии с теорией Менделя генотип потомка определяется следующим образом. Каждая из двух *гамет* (половых клеток) каждого из двух *родителей* переносит один *ген* g или G из составляющих генотип этого родителя. Для участия в образовании *зиготы* (оплодотворенной клетки) у каждого из родителей наугад выбирается одна из га-

мет. Гены, принесенные этими гаметами, и образуют генотип потомка.

В частности, эта модель хорошо объясняет результаты опыта по скрещиванию растений с белыми и красными цветами, о котором говорилось в пункте 2 из § 1.

В задаче о красных, белых и розовых урнах эти урны играют роль генотипов, а шары в них — генов. Нумерация шаров и разбиение розовых урн на два типа не имеют, по-видимому, генетического смысла и введены для симметричности схемы.

Пример 1. *Имеются 130 растений с красными цветами, 290 гибридов с розовыми и 140 растений с белыми цветами. Какова вероятность того, что при скрещивании наугад выбираемых растений получается растение с белыми цветами?*

Решение. Предположим, что наследование окраски цветов описывается рассматриваемой моделью. Дело сводится к задаче о красных, белых и розовых урнах при $a=130$, $b=145$, $c=140$. Искомая вероятность равна

$$\left(\frac{b+c}{a+2b+c}\right)^2=\left(\frac{295}{560}\right)^2\approx 0{,}28.$$

Пример 2. *Имеются 130 растений с красными цветами и 290 гибридов с розовыми цветами. Какова вероятность того, что при скрещивании наугад выбираемых растений получается растение с белыми цветами?*

Решение. По-прежнему предполагаем, что наследование окраски цветов растений описывается рассматриваемой моделью. Дело сводится к задаче о красных, белых и розовых урнах при $a=130$, $b=145$, $c=0$. Искомая вероятность равна

$$\left(\frac{b}{a+2b}\right)^2=\left(\frac{145}{420}\right)^2\approx 0{,}12.$$

Замечание. Этот пример поясняет влияние *селекции.* Если растения с белыми цветами исключить из процесса размножения, то вероятность появления потомка с белыми цветами уменьшается, а вероятность появления потомка с красными цветами увеличивается:

$$\left(\frac{a+b}{a+2b+c}\right)^2=\left(\frac{275}{560}\right)^2\approx 0{,}24;\quad \left(\frac{a+b}{a+2b}\right)^2=\left(\frac{275}{420}\right)^2\approx 0{,}43.$$

Вероятности появления гибридов с розовыми цветами приблизительно равны соответственно 0,48 и 0,45.

Пример 3. *Известно, что при скрещивании наугад выбираемых из 130 растений с красными цветами и 290 растений с розовыми появляется потомок с красными цветами. Какова вероятность того, что его родители красные цветы?*

Решение. В модели задачи с красной, белой и розовой урнами при $a=130$, $b=145$, $c=0$ дело сводится к вычислению вероятности события $B\begin{pmatrix}11\\11\end{pmatrix}$ при условии $A\begin{pmatrix}1\\1\end{pmatrix}$. По формуле Байеса

$$P_{A\binom{1}{1}}\left[B\begin{pmatrix}11\\11\end{pmatrix}\right]=P\left[B\begin{pmatrix}11\\11\end{pmatrix}\right]P_{B\binom{11}{11}}\left[A\begin{pmatrix}1\\1\end{pmatrix}\right]\Big/P\left[A\begin{pmatrix}1\\1\end{pmatrix}\right]=$$

$$=\left(\frac{a}{a+2b}\right)^2\Big/\left(\frac{a+b}{a+2b}\right)^2=\left(\frac{a}{a+b}\right)^2=\left(\frac{130}{275}\right)^2\approx 0{,}22.$$

Глава 2

СЛУЧАЙНЫЕ ПЕРЕМЕННЫЕ

В этой главе рассматриваются некоторые задачи, связанные со случайными переменными для конечной вероятностной модели.

§ 1. СРЕДНЕЕ И ДИСПЕРСИЯ

Рассмотрим конечную вероятностную модель, определяемую *множеством исходов* U и *элементарной вероятностью* p.

Случайной переменной называется каждая числовая функция f на множестве исходов U.

Случайные переменные f и g называются *независимыми*, если они принимают значения независимо друг от друга: для каждых чисел x и y событие $A(x)=\{u:f(u)=x\}$ (f принимает значение x) и событие $B(y)=\{u:g(u)=y\}$ (g принимает значение y) независимы. Аналогично определяется независимость нескольких случайных переменных.

1.1. Среднее

Важнейшей числовой характеристикой случайной переменной f является ее *среднее* $E(f)$, определяемое как *сумма* произведений вероятностей $p(u)$ исходов u на соответствующие значения $f(u)$ случайной переменной f:

$$E(f)=\Sigma f(u)p(u).$$

При вычислении среднего удобно использовать следующие правила:

Правило постоянной. $E(c)=c$;

Правило сложения. $E(f+g)=E(f)+E(g)$;

Правило умножения на число. $E(cf)=cE(f)$;

Правило неравенства. $E(f)\leqslant E(g)\qquad(f\leqslant g)$.

Если случайные переменные f и g *независимы*, то для них верно

Правило умножения. $E(f \cdot g) = E(f) \cdot E(g)$.

Из определения среднего следует, что оно равно также сумме произведений значений x случайной переменной f на вероятности $q(x)$ этих значений:

$$E(f) = \Sigma x q(x), \quad (q(x) = P\{u : f(u) = x\}).$$

Это равенство часто сокращает вычисление среднего.

Рассмотрим несколько примеров вычисления среднего.

Задача 1. *Игральная кость подбрасывается два раза. Какова средняя сумма очков?*

Решение. В соответствии с задачей 1 § 1 главы 1 эта сумма описывается суммой f случайных переменных f_1 и f_2, выражающих соответственно число очков при первом и втором подбрасываниях:

$$f_1(u) = u_1, \ f_2(u) = u_2, \ f(u) = u_1 + u_2 \ (u = u_1 u_2).$$

По правилу сложения отсюда следует, что

$$E(f) = E(f_1) + E(f_2) = 7.$$

Замечание. Задача 4 § 1 главы 1 показывает, что эта средняя сумма является и *наиболее вероятной.* Таблица вероятностей значений для суммы очков, полученная при решении этой задачи, позволяет вычислить среднюю сумму очков непосредственно.

Задача 2. *Из урны с шарами 1, 2, 3, 4, 5 два раза наугад вынимается шар и возвращается в урну. Какова средняя разность номеров?*

Решение. В соответствии с задачей 2 § 1 главы 1 эта разность описывается разностью f случайных переменных f_1 и f_2, выражающих соответственно номер шара при первом и втором извлечениях:

$$f_1(u) = u_1, \ f_2(u) = u_2, \ f(u) = u_1 - u_2 \ (u = u_1 u_2).$$

Используя указанное для вычисления среднего равенство, получаем:

$$E(f_1) = E(f_2) = 1\frac{1}{5} + 2\frac{1}{5} + 3\frac{1}{5} + 4\frac{1}{5} + 5\frac{1}{5} = 3.$$

По правилам сложения и умножения на число отсюда следует, что

$$E(f) = E(f_1) - E(f_2) = 0.$$

Задача 3. *Из урны с шарами 1, 2, 3, 4, 5 два раза наугад вынимается шар и не возвращается в урну. Какова средняя разность номеров?*

Решение. В соответствии с задачей 3 § 1 главы 1 эта разность описывается аналогично задаче 2. По-прежнему

$$E(f_1) = (1+2+3+4+5)\,1/5 = 3.$$

По определению

$$E(f_2)=\sum_u f_2(u)\frac{1}{20}=\frac{1}{20}[(15-1)+(15-2)+(15-3)+$$
$$+(15-4)+(15-5)]=3.$$

Снова

$$E(f)=E(f_1)-E(f_2)=0.$$

Замечание. Среднее $E(f_2)$ можно вычислить так же, как и $E(f_1)$, если, используя формулу полной вероятности, доказать, что f_2 принимает каждое значение $x=1, 2, 3, 4, 5$ с вероятностью

$$P\{u:f_2(u)=x\}=4\cdot 1/5\cdot 1/4=1/5.$$

Задача 4. *Вычислительная машина производит $n=10^6$ одинаковых и независимых операций, в каждой из которых с вероятностью $a=10^{-6}$ происходит ошибка. Каково среднее число ошибок?*

Решение. Оно равно среднему $E(s)$ числа успехов s для модели Бернулли $n=10^6$ испытаний с вероятностью успеха $a=10^{-6}$. Общее число успехов s равно сумме чисел успехов s_j при j-м испытании $(1\leqslant j\leqslant n)$. Используя правило сложения, получаем:

$$E(s)=\sum_{1\leqslant j\leqslant n} E(s_j)=\sum_{1\leqslant j\leqslant n}(1\cdot a+0(1-a))=na=1.$$

Диффузия газов. Рассмотрим 2 литровые сосуда, один из которых содержит газ A, а другой — газ B, находящиеся под атмосферным давлением и при одинаковой температуре. Если сосуды соединить трубкой, то через некоторое время в каждом из них окажется одинаковая смесь C газов A и B.

Можно так объяснить это явление. Каждый из газов представляет собой множество $n=10^{22}$ хаотически двигающихся молекул. Поведение каждого из газов после соединения сосудов описывается моделью Бернулли $n=10^{22}$ испытаний с вероятностью успеха $a=1/2$ (успехом считается переход молекулы в другой сосуд). Вследствие хаотичности движения молекул можно считать, что доля диффундировавших молекул описывается частотой успеха $n^{-1}\cdot s$. А установившееся равновесие связано с равенством средней частоты успеха его вероятности 1/2.

Задача 5. *Какова средняя доля диффундирующих молекул?*

Решение. Используя результат задачи 4 и правило умножения на число, получаем:

$$E(n^{-1}s)=n^{-1}E(s)=n^{-1}na=a=1/2.$$

Замечание. В § 1 главы 1 было показано, что наиболее вероятное значение для числа успехов

$$\hat{m}=[(n+1)a]=2^{-1}\cdot 10^{22}.$$

Следовательно, наиболее вероятное значение для частоты успеха равно

$$n^{-1}\hat{m}=2^{-1}.$$

Таким образом, установление равновесия является и наиболее вероятным событием. В то же время вероятность этого события, т. е. вероятность того, что диффундирует ровно половина молекул, меньше 10^{-10}.

Задача 6. *Симметричная монета подбрасывается n раз. Если она впервые падает гербом вверх при j-м подбрасывании, то игрок выигрывает 2^j рублей ($j=1, \ldots, n$). Если при всех n подбрасываниях монета падает цифрой вверх, то игрок проигрывает $n \cdot 2^n$ рублей. Каков средний выигрыш игрока?*

Решение. Этот выигрыш описывается случайной переменной f на множестве исходов модели Бернулли n испытаний с вероятностью успеха $a=1/2$, определяемой следующим образом: $f(u)=2^j$, если в строке u первая по порядку единица расположена на j-м месте; кроме того, $f(0\ldots0)=-n\cdot 2^n$. Задача сводится к вычислению среднего $E(f)$ случайной переменной f.

Она принимает значение $x=2^j$ с вероятностью $q(x)=2^{-j}$. Это следует из того, что число строк u длины n из 1 и 0, у которых первая по порядку единица расположена на j-м месте, равно числу всех строк длины $n-j$ из 1 и 0, т. е. числу 2^{n-j}. А вероятность каждого исхода равна 2^{-n}. Значит,

$$q(x)=2^{n-j}2^{-n}=2^{-j} \quad (x=2^j).$$

Используя указанное для вычисления среднего равенство, получаем:

$$E(f)=\sum_{1\leqslant j\leqslant n} 2^j 2^{-j} - n\cdot 2^n \cdot 2^{-n} = n-n=0.$$

Задача 7. *Каково среднее число мест, на которых в двух идентичных хорошо тасованных колодах находятся одинаковые карты?*

Решение. Каждое взаимное расположение n карт в колодах можно описать перестановкой множества из n элементов. Так как таких перестановок ровно $n!$, то в качестве модели для рассматриваемой задачи о совпадениях удобно взять модель Лапласа $n!$ равновероятных исходов, которыми являются перестановки номеров $1, \ldots, n$.

Для каждого номера $j=1, \ldots, n$ рассмотрим случайную переменную f_j, значение $f_j(u)$ которой равно 1, если перестановка n оставляет номер j на месте $(u(j)=j)$, и равно 0 в противном случае $(u(j)\neq j)$. Число совпадений в рассматриваемых колодах описывает сумма f случайных переменных f_j $(j=1, \ldots, n)$.

Число перестановок u, оставляющих номер j *на месте*, равно числу всех перестановок остальных $n-1$ номеров, т. е. числу $(n-1)!$. Следовательно,

$$E(f_j)=1\,\frac{(n-1)!}{n!}+0\left(1-\frac{(n-1)!}{n}\right)=1/n \quad (j=1, \ldots, n).$$

По правилу сложения отсюда вытекает, что

$$E(f)=\sum_{1\le j\le n} E(f_j)=n\frac{1}{n}=1.$$

Замечание. Таким образом, среднее число совпадений равно единице и не зависит от числа карт в колоде. Этот результат вряд ли можно было ожидать.

Задача 8. *Из урны, содержащей 1 красный и $n-1$ белых шаров, вынимаются по одному шары (без возвращения). Каково среднее число белых шаров, вынутых перед красным?*

Решение. Как и в задаче о красных шарах из § 4 главы 1, предположим, что шары отмечены номерами 1, …, n, причем красный шар имеет номер 1. Тогда каждый выбор n шаров из урны можно описать перестановкой u множества номеров $\{1, \ldots, \ldots, n\}$. В качестве модели для решения рассматриваемой задачи возьмем, как и в предыдущей, модель Лапласа $n!$ равновероятных исходов, которыми являются такие перестановки u.

Рассмотрим случайную переменную f, значение которой $f(u)$ для каждой перестановки $f(u)$ равно номеру $u(1)$ места, на которое переставляется номер 1, уменьшенному на единицу:

$$f(u)=u(1)-1.$$

Для каждого номера $x=0, \ldots, n-1$ число перестановок u, для которых $u(1)=x+1$, равно числу подстановок вместо $\{2, \ldots, n\}$ множества $\{1, \ldots, n\}-\{x+1\}$, т. е. числу $(n-1)!$. Следовательно, вероятность $q(x)$ того, что случайная переменная f примет значение x, выражается равенствами

$$q(x)=\frac{(n-1)!}{n!}=1/n \qquad (x=0, \ldots, n-1).$$

Поэтому

$$E(f)=\sum_{0\leqslant x\leqslant n} x\frac{1}{n}=\frac{1}{n}\cdot\frac{n(n-1)}{2}=\frac{n-1}{2}.$$

2.1. Дисперсия

Среднее является очень грубой характеристикой случайной переменной: ее значения могут сильно отклоняться от него. Поэтому для более точного описания случайной переменной f используется также *дисперсия* $D(f)$, определяемая как *среднее квадрата отклонения* f *от* $E(f)$:

$$D(f)=E[(f-E(f))^2]=\Sigma(f(u)-E(f))^2p(u).$$

При вычислении дисперсии удобно использовать равенства

$$D(f)=\Sigma(x-E(f))^2q(x),$$
$$D(f)=E(f^2)-E^2(f)$$

и следующие правила:

Правило постоянной. $D(c)=0$;

Правило умножения на число. $D(cf)=c^2D(f)$;

Правило прибавления постоянной. $D(f+c)=D(f)$.

Если случайные переменные f и g независимы, то для них верно

Правило сложения. $D(f+g)=D(f)+D(g)$.

Рассмотрим несколько примеров вычисления дисперсии. Используем случайные переменные, описанные в предыдущем пункте, и вычисленные для них средние.

Задача 1. Используя указанные для вычисления дисперсии равенства, получаем:

$$D(f_1)=D(f_2)=\sum_{1\leqslant x\leqslant 6}(x-3{,}5)^2\frac{1}{6}\approx 2{,}91.$$

Вследствие *независимости* случайных переменных f_1 и f_2 по правилу сложения отсюда вытекает, что

$$D(f)=D(f_1)+D(f_2)\approx 5{,}82.$$

Задача 2. Аналогично:

$$D(f_1)=D(f_2)=\sum_{1\leqslant x\leqslant 5}(x-3)^2\frac{1}{5}=\frac{5}{3},$$

$$D(f)=D(f_1)+D(f_2)=\frac{10}{3}\approx 3{,}3.$$

Задача 3. В этой задаче случайные переменные f_1 и f_2 *зависимы*, и поэтому правило сложения для дисперсий к ним не применимо. Используем формулу

$$D(f)=E[(f-E(f))^2].$$

Так как $E(f)=0$, то

$$D(f)=E(f^2).$$

Используя правила для вычисления средних, получаем отсюда

$$D(f)=E((f_1-f_2)^2)=E\left(f_1^2\right)+E\left(f_2^2\right)-2E(f_1f_2).$$

Пользуясь указанным для вычисления дисперсии равенством, убеждаемся в том, что

$$E\left(f_1^2\right)=E\left(f_2^2\right)=\sum_{1\leqslant x\leqslant 5}x^2\frac{1}{5}=11.$$

Из определений вытекает, что

$$E(f_1f_2)=\sum_{x\neq y}xy\frac{1}{20}=\frac{1}{20}\left[\left(\sum_{1\leqslant x\leqslant 5}x\right)\left(\sum_{1\leqslant y\leqslant 5}y\right)-\sum_{1\leqslant z\leqslant 5}z^2\right]=$$

$$=\frac{1}{20}(15\cdot 15-55)=8{,}5.$$

Следовательно,

$$D(f)=11+11-17=5.$$

Замечание. Дисперсия разности номеров вынимаемых из урны шаров для выбора без возвращения больше, чем для выбора с возвращением. Это объясняется тем, что при выборе без возвращения исключены равные нулю значения разности номеров.

Задача 4. Используя указанное для вычисления дисперсии равенство, получаем:

$$D(s_j) = (1-a)^2 a + (0-a)^2(1-a) = a(1-a) = 10^{-6}(1-10^{-6}) \quad (1 \leqslant j \leqslant n).$$

Так как случайные переменные s_j *независимы*, то, по правилу сложения,

$$D(s) = \sum_{1 \leqslant j \leqslant n} D(s_j) = na(1-a) = 1-10^{-6}.$$

Задача 5. Аналогично, используя правило умножения на число, получаем:

$$D(n^{-1}s) = n^{-2}D(s) = n^{-1}a(1-a) = 4^{-1}n^{-1}.$$

Задача 6. Используя указанное для вычисления дисперсии равенство, получаем:

$$D(f) = \sum_{1 \leqslant j \leqslant n} (2^j - 0)^2\, 2^{-j} + (n^2 2^{2n} - 0)\, 2^{-n} = \sum_{1 \leqslant j \leqslant n} 2^j + n^2 2^n = $$
$$= (n^2+2)\,2^n - 2.$$

Для достаточно больших номеров n дисперсия произвольно велика.

Задача 7. Вычислим предварительно средние произведений $f_j f_k$:

$$E(f_j^2) = P\{u : f_j^2(u) = 1\} = P\{u : f_j(u) = 1\} = \frac{(n-1)!}{n!} = \frac{1}{n} \quad (1 \leqslant j \leqslant n),$$

$$E(f_j f_k) = P\{u : f_j(u) = f_k(u) = 1\} = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)} \quad (1 \leqslant j \neq k \leqslant n).$$

Используя эти равенства и правила вычисления среднего, получаем:

$$E(f_j^2) = \sum_{1 \leqslant j \leqslant n} E(f_j^2) + \sum_{1 \leqslant j \neq k \leqslant n} E(f_j f_k) = n\frac{1}{n} + n(n-1)\frac{1}{n(n-1)} = 2.$$

Отсюда и из формулы квадратов для вычисления дисперсии следует, что

$$D(f) = E(f^2) - E^2(f) = 2 - 1 = 1.$$

Как и среднее, дисперсия числа совпадений равна единице и не зависит от числа карт в колодах.

Задача 8. Используя равенство для суммы квадратов из главы 2 части II раздела I, получаем:

$$D(f) = E(f^2) - E^2(f) = \sum_{0 \leqslant x \leqslant n} x^2 \frac{1}{n} - \frac{(n-1)^2}{4} =$$
$$= \left(\frac{1}{3}n^3 - \frac{1}{2}n^2 + \frac{1}{6}n\right)\frac{1}{n} - \frac{(n-1)^2}{4} = \frac{n^2-1}{12}.$$

§ 2. ЗАКОН БОЛЬШИХ ЧИСЕЛ

Этот закон для *попарно независимых* и *одинаково распределенных* случайных переменных f_j со средним и дисперсией

$$E(f_j)=a,\ D(f_j)=b^2\ \ (b>0,\ 1\leqslant j\leqslant n)$$

на множестве исходов U конечной вероятностной модели $(U, \mathscr{A}; p, P)$ выражает

Теорема Чебышева.

$$P\left\{u:\left|n^{-1}\sum_{1\leqslant j\leqslant n} f_j(u)-a\right|\geqslant\varepsilon\right\}\leqslant\varepsilon^{-2}n^{-1}b^2\quad(\varepsilon>0).$$

Она оценивает вероятность *значительного* отклонения *арифметического среднего* случайных переменных f_j от их *среднего* a. Для достаточно больших чисел n это отклонение *практически невозможно*.

Из теоремы Чебышева следует

Теорема Бернулли.

$$P\left\{u:\left|n^{-1}\sum_{1\leqslant j\leqslant n} s_j(u)-a\right|\geqslant\varepsilon\right\}\leqslant\varepsilon^{-2}\,n^{-1}a(1-a)\leqslant\varepsilon^{-2}n^{-1}4^{-1}\qquad(\varepsilon>0).$$

Она оценивает вероятность *значительного* отклонения *частоты* успеха от его *вероятности*. Для достаточно больших чисел n это отклонение *практически невозможно*.

Теоремы Чебышева и Бернулли дают очень грубые оценки рассматриваемых вероятностей. В § 5 главы 3 части II рассматриваются более точные неравенства.

1.2. Примеры

Рассмотрим несколько примеров использования закона больших чисел. Положим:

$$B(n,\varepsilon)=\left\{u:\left|n^{-1}\sum_{1\leqslant j\leqslant n} f_j(u)-a\right|\geqslant\varepsilon\right\}.$$

1.1.2. Пример 1

Для задачи 4 из § 1 о вычислительной машине по теореме Бернулли получаем при $\varepsilon=2^{-1}\cdot10^{-2}$

$$P(B(n,\varepsilon))\leqslant4\cdot10^4\cdot10^{-6}\cdot4^{-1}=10^{-2}.$$

Если считать значительными отклонения больше, чем на $\varepsilon=10^{-2}$, и считать практически невозможными события, вероятности которых меньше $\alpha=10^{-2}$, то из этого неравенства следует, что значительное отклонение частоты ошибки от ее вероятности $a=10^{-6}$ практически невозможно.

2.1.2. Пример 2

Для задачи 5 из § 1 о диффузии газов по теореме Бернулли получаем при $\varepsilon=10^{-8}$

$$P(B(n, \varepsilon)) \leqslant 10^{16} \cdot 10^{-22} \cdot 4^{-1} < 10^{-6}.$$

Если считать значительными отклонения больше, чем на $\varepsilon = 10^{-8}$, и считать практически невозможными события, вероятности которых меньше $\alpha=10^{-6}$, то из этого неравенства следует, что значительное отклонение доли диффундирующих молекул от $a=1/2$ практически невозможно. Именно вследствие закона больших чисел можно считать, что средняя доля диффундирующих молекул описывает рассматриваемую смесь газов.

3.1.2. Пример 3

Для задачи 7 из § 1 о числе совпадений по теореме Чебышева получаем при $\varepsilon=n^{-1}$:

$$P(B(n, \varepsilon)) \leqslant n^2 n^{-1} n^{-1}(1-n^{-1}) = 1-n^{-1}.$$

Заметим, что в рассматриваемом случае $a=\varepsilon=n^{-1}$, и поэтому

$$B(n, \varepsilon) = \left\{u : \left| n^{-1} \sum_{1 \leqslant j \leqslant n} f_j(u) - n^{-1} \right| \geqslant n^{-1} \right\} = \left\{u : \left| \sum_{1 \leqslant j \leqslant n} f_j(u) - 1 \right| \geqslant 1 \right\} = \left\{u : \sum_{1 \leqslant j \leqslant n} f_j(u) \neq 1 \right\}.$$

Следовательно,

$$P\left\{u : \sum_{1 \leqslant j \leqslant n} f_j(u) \neq 1 \right\} \leqslant 1 - n^{-1}.$$

$$P\left\{u : \sum_{1 \leqslant j \leqslant n} f_j(u) = 1 \right\} \geqslant n^{-1},$$

Полученные оценки плохие.

Теорема Бернулли позволяет решить две часто встречающиеся задачи: 1) *о проверке гипотезы* и 2) *об оценке вероятности.*

2.2. Проверка гипотезы

Рассмотрим последовательность n испытаний, описываемую моделью Бернулли с неизвестной вероятностью успеха. Сделаем предположение, что эта вероятность равна a $(0 \leqslant a \leqslant 1)$. При реализации рассматриваемой последовательности частота успехов

оказывается равной b ($0 \leqslant b \leqslant 1$). Согласуется ли сделанное предположение с результатом эксперимента?

Представляется разумным использовать следующий принцип: если при реализации произошло событие, которое при сделанном предположении является практически невозможным, то предположение не согласуется с результатом эксперимента.

Рассмотрим модель Бернулли с параметрами n и a. Условимся считать практически невозможными события, вероятность которых меньше α ($0<\alpha<1$). Если сделанное предположение о том, что вероятность успеха равна a, соответствует действительности, то для вероятности $P(B)$ события

$$B = \{u: |(1/n)s(u) - a| \geqslant \Delta\},$$

где $\Delta = |b-a|$, верно неравенство

$$P(B) \leqslant a(1-a)/n\Delta^2.$$

Если

$$\Delta \geqslant \Delta_0 = \sqrt{\frac{a(1-a)}{n\alpha}},$$

то

$$a(1-a)/n\Delta^2 \leqslant \alpha$$

и

$$P(B) \leqslant \alpha,$$

т. е. при сделанном предположении *отклонение частоты успеха от вероятности успеха больше, чем на величину* Δ_0, *является практически невозможным событием.*

Таким образом, проверка гипотезы о равенстве вероятности успеха числу a для модели Бернулли n испытаний заключается в следующем.

1. Выбирается *оценка практической невозможности* α и определяется *допустимое отклонение* Δ_0 значения частоты успеха от вероятности успеха a:

$$\Delta_0 = \sqrt{\frac{a(1-a)}{n\alpha}};$$

2. Подсчитывается полученное в результате эксперимента *значение b частоты* успеха и определяется *отклонение*

$$\Delta = |b-a|;$$

3. Если *полученное отклонение* Δ значения частоты успеха от вероятности успеха *больше допустимого* Δ_0:

$$\Delta \geqslant \Delta_0,$$

то гипотеза *не согласуется* с имеющимися данными.

Замечание. Если $\Delta < \Delta_0$, то отсюда *нельзя* сделать вывод о том, что гипотеза согласуется с имеющимися данными. Из-за гру-

бости использованной оценки событие B может оказаться практически невозможным и при $\Delta < \Delta_0$.

Рассмотрим несколько примеров.

1.2.2. Задача о жулике

Яркий пример вопиющего несоответствия между предположением и действительностью описан в следующем старинном анекдоте: «Однажды в Неаполе преподобный Галиани увидел человека из Базиликаты, который, встряхивая три игральные кости в чашке, держал пари, что выбросит три шестерки. Вы скажете, такая удача возможна. Однако человеку из Базиликаты удавалось это и во второй раз, и пари повторялось. Он клал кости в чашку 3, 4, 5 раз и каждый раз выбрасывал три шестерки. «Черт возьми! — воскликнул преподобный.— Кости налиты свинцом!» Так оно и было. Но почему преподобный воспользовался нечестивым выражением?» Преподобный Галиани выругался, очевидно, решив, что человек из Базиликаты — жулик. Соответствует ли действительности предположение о том, что человек из Базиликаты — честный?

Процесс 5-кратного честного бросания трех честных костей можно описать моделью Бернулли $n=5$ испытаний с вероятностью успеха $a=1/6^3$ (успехом считается выпадение трех шестерок). Предположим, что человек из Базиликаты и его кости — честные. Задача, таким образом, сводится к проверке гипотезы о равенстве вероятности успеха данному значению. Используем построенную для этой проверки схему.

1. Условимся считать практически невозможными события, вероятности которых меньше $\alpha=0{,}001$. В этом случае допустимое отклонение $\Delta_0=[6^{-3}(1-6^{-3})/(5\cdot 10^{-3})]^{1/2}$.

2. Полученное отклонение $\Delta=1-6^{-3}$.

3. Нетрудно проверить, что $\Delta \geqslant \Delta_0$.

В самом деле, это неравенство эквивалентно каждому из следующих неравенств, последнее из которых очевидно: $(1-6^{-3})^2 \geqslant 6^{-3}(1-6^{-3})/(5\cdot 10^{-3})$; $1-6^{-3} \geqslant 6^{-3}/(5\cdot 10^{-3})$; $5(6^3-1) \geqslant 10^3$; $1075 \geqslant 1000$.

Таким образом, полученное отклонение значения частоты успеха от вероятности успеха больше допустимого. Предположение, что человек из Базиликаты — честный человек, *не согласуется* с имеющимися данными. Преподобный Галиани был прав.

2.2.2. Статистический контроль производственного процесса

Условимся работу автомата, производящего гвозди, описывать схемой Бернулли с вероятностью успеха $a=1/4$ (успехом считается изготовление негодного гвоздя). При проверке партии из $n=10\,000$ гвоздей доля негодных оказалась равной $b=5/12$.

Можно ли считать, что это свидетельствует о неполадках в работе автомата?

Задача сводится к проверке гипотезы о равенстве вероятности успеха данному значению. Используем построенную для этой проверки схему.

1. Условимся считать практически невозможными события, вероятности которых меньше $\alpha=0{,}001$. В этом случае допустимое отклонение $\Delta_0=[1/4\cdot 3/4)/(10^4\cdot 10^{-3})]^{1/2}<[1/4\cdot 10]^{1/2}<1/6$.

2. Полученное отклонение $\Delta=|5/12-1/4|=1/6$.

3. Таким образом, полученное отклонение значения частоты успеха от вероятности успеха больше допустимого: $\Delta\geqslant\Delta_0$.

Поэтому можно считать, что полученные данные свидетельствуют о неполадках в работе автомата.

3.2.2. Рак легких

Условимся считать, что процесс выявления заболевания раком легких при обследовании группы некурящих людей описывается моделью Бернулли с вероятностью успеха $a=10^{-4}$ (успехом считается то обстоятельство, что человек болен раком легких). При обследовании группы $n=10^5$ курящих людей доля больных раком оказалась равной $b=11\cdot 10^{-4}$. Можно ли считать, что эти данные подтверждают связь между курением и раком легких?

Задача сводится к проверке гипотезы о равенстве вероятности успеха по данному значению. Используем построенную для этой проверки схему.

1. Условимся считать практически невозможными события, вероятности которых *меньше* $\alpha=0{,}001$. В этом случае допустимое отклонение $\Delta_0=[10^{-4}(1-10^{-4})/(10^5\cdot 10^{-3})]^{1/2}<10^{-3}$.

2. Полученное отклонение $\Delta=11\cdot 10^{-4}-10^{-4}=10^{-3}$.

3. Таким образом, полученное отклонение значения частоты успеха от вероятности успеха больше допустимого: $\Delta\geqslant\Delta_0$.

Можно считать, что имеющиеся данные подтверждают существование связи между курением и раком легких.

4.2.2. Пол ребенка

Условимся считать, что процесс рождения мальчиков и девочек описывается схемой Бернулли с неизвестной вероятностью успеха (успехом будем считать рождение мальчика). Представляется правдоподобным предположение, что рождение мальчика и рождение девочки равновероятны. По имеющимся данным, в Швейцарии с 1871 по 1900 год родилось $n=2\,644\,757$ детей. Среди них было $m=1\,359\,671$ мальчик. Согласуется ли предположение о равенстве вероятностей рождения мальчика и рождения девочки с этими данными?

Задача сводится к проверке гипотезы о том, что вероятность успеха $a=1/2$. Используем построенную для проверки подобных гипотез схему.

1. Условимся считать практически невозможными события, вероятности которых меньше $\alpha=0{,}001$. В этом случае допустимое отклонение $\Delta_0=[(1/4)/(2\,644\,757\cdot10^{-3})]^{1/2}<[4\cdot25\cdot10^5\cdot10^{-3}]^{-1/2}=$ $=0{,}01$.

2. Значение частоты успеха для рассматриваемого случая $b\geqslant0{,}5141$. Следовательно, полученное отклонение $\Delta\geqslant0{,}5141-$ $-0{,}5000=0{,}0141$.

3. Таким образом, полученное отклонение значения частоты успеха от вероятности успеха больше допустимого: $\Delta\geqslant\Delta_0$.

Предположение, что рождение мальчика и рождение девочки равновероятны, не согласуется с имеющимися данными. (В демографии принято считать, что вероятность рождения мальчика равна 0,515).

5.2.2. Телепатия

В контрольном опыте по проверке существования телепатической связи (Москва — Керчь), проведенном 10 мая 1968 года, из 10 телепатически переданных образов предметов правильно не был принят ни один. Предположим, что прием описывается моделью Бернулли для $n=10$ испытаний. Согласуется ли результат опыта с предположением, что вероятность правильного приема образов равна $a=0{,}999$?

Задача сводится к проверке гипотезы о том, что вероятность успеха $a=0{,}999$. Используем построенную для проверки схему.

1. Условимся считать практически невозможными события, вероятности которых меньше $\alpha=0{,}001$. В этом случае допустимое отклонение $\Delta_0=[(1-10^{-3})\cdot10^{-3}/(10\cdot10^{-3})]^{1/2}<0{,}33$.

2. Полученное отклонение $\Delta=|0-0{,}999|=0{,}999$.

3. Таким образом, полученное отклонение в 3 раза превышает допустимое: $\Delta>3\Delta_0$.

Можно считать, что гипотеза о практически достоверном приеме телепатически передаваемого образа предмета контрольным опытом не подтвердилась. (В то же время, разумеется, проделанный расчет нельзя считать доказательством отсутствия телепатической связи вообще).

Замечание. Во всех рассмотренных примерах практически невозможные события были определены как события, вероятности которых меньше $\alpha=0{,}001$. Выбор значения 0,001 произволен. Обычно α определяется в зависимости от конкретных обстоятельств. Уменьшение α соответствует менее строгому подходу к проверке гипотезы: допустимое отклонение увеличивается и вывод о несогласованности с результатом эксперимента делается реже.

3.2. Оценка вероятности

Рассмотрим вновь последовательность n испытаний, описываемую моделью Бернулли с неизвестной вероятностью успеха. В качестве *оценки для вероятности* успеха используем *частоту*

успеха. Какое число испытаний *практически достоверно* обеспечивает *данную точность* оценки?

Условимся считать практически достоверными события, вероятность которых больше $1-\alpha$ $(0<\alpha<1)$. Вместо практически достоверно будем говорить также с вероятностью, большей $1-\alpha$. Точность оценки $(1/n)s$ зададим числом ε $(0<\varepsilon<1)$. Оцениваемую вероятность обозначим буквой a $(0\leqslant a\leqslant 1)$.

Задача сводится к определению числа n_0 такого, что для каждого натурального числа $n\geqslant n_0$ и каждого числа a $(0\leqslant a\leqslant 1)$ в модели Бернулли с параметрами n и a было верно неравенство

$$P\{u: |(1/n)s-a|<\varepsilon\}\geqslant 1-\alpha.$$

По теореме Бернулли, для этого достаточно, чтобы $n_0=1/4\alpha\varepsilon^2$.

Таким образом, если число испытаний n в схеме Бернулли *больше* числа

$$n_0=1/4\alpha\,\varepsilon^2,$$

то *с вероятностью, большей чем $1-\alpha$, частота успехов $(1/n)s$ оценивает вероятность успеха a с точностью ε.*

Рассмотрим несколько примеров.

1.3.2. Статистический контроль качества продукции

Вернемся к автомату, производящему гвозди. Предположим, что он заменен новым, работу которого также можно описать моделью Бернулли, но уже с неизвестной вероятностью успеха a (успехом по-прежнему считается изготовление негодного гвоздя). Для оценки этой вероятности производится пробная партия из n гвоздей. В качестве оценки для a берется *доля* негодных гвоздей пробной партии. Требуется определить число n гвоздей для пробной партии, при котором эта оценка практически достоверно имела бы точность $\varepsilon=0{,}1$. Практически достоверные события определим как события, вероятности которых больше $1-\alpha=0{,}999$ $(\alpha=0{,}001)$.

Задача сводится к определению числа испытаний, при котором использование частоты успеха практически достоверно обеспечивает заданную точность. Используем полученное для этого числа неравенство. В рассматриваемом случае $\alpha=0{,}001$ и $\varepsilon=0{,}1$, следовательно, $n\geqslant n_0=2500$.

2.3.2. Оценка доли курящих

Неизвестная доля a жителей города курит. Для определения доли курящих жителей предполагается провести ряд независимых наблюдений, при которых каждый раз все жители будут иметь равную возможность стать объектом наблюдения. Можно поэтому считать, что процесс наблюдения описывается схемой n испы-

таний Бернулли с вероятностью успеха a (успехом считается то, что наблюдаемый житель курит). В качестве оценки для доли a курящих жителей города предполагается использовать частоту s/n курящих среди наблюдавшихся жителей. Требуется, чтобы оценка практически достоверно имела точность $\varepsilon=0{,}005$. Практически достоверными считаются события, вероятности которых больше $1-\alpha=0{,}95$ $(\alpha=0{,}05)$. Сколько нужно провести наблюдений?

Задача сводится к определению числа испытаний, при котором использование частоты успеха практически достоверно обеспечивает заданную точность. Используем полученное для этого числа неравенство. В рассматриваемом случае $\alpha=0{,}05$ и $\varepsilon=0{,}005$. Следовательно, нужно провести $n\geqslant n_0=200\,000$ наблюдений.

3.3.2. Эффективность лечения

Испытывается новый метод лечения некоторой болезни. О его эффективности предполагается судить по доле выздоровевших среди подопытных зараженных кроликов, подвергнутых лечению данным методом. Требуется, чтобы оценка эффективности была достаточно точной и надежной. Сколько нужно кроликов для такого опыта?

Будем предполагать, что процесс выздоровления или смерти подопытных кроликов описывается моделью Бернулли n испытаний с неизвестной вероятностью успеха (выздоровление кролика) a. Практически достоверные события определим как события, вероятность которых *больше* $1-\alpha=0{,}95$ $(\alpha=0{,}05)$. *Точность* оценки зададим числом $\varepsilon=0{,}25$. Задача о числе подопытных кроликов сводится к определению числа n испытаний, при котором оценка вероятности успеха a с помощью частоты успеха s/n практически достоверно имеет точность ε. Применяя полученное для этого числа неравенство, находим, что $n\geqslant n_0=80$.

Если требовать при той же практической достоверности точность $\varepsilon=0{,}05$, то будет $n\geqslant n_0=2\,000$.

Если при этой точности оценивать практическую достоверность числом $1-\alpha=0{,}99$ $(\alpha=0{,}01)$, получим $n\geqslant n_0=10\,000$.

Замечание. Полученные с помощью неравенства Чебышева оценки для отклонения частоты успеха от вероятности успеха и для числа испытаний являются довольно грубыми. Более развитая теория дает более точные оценки. В частности, можно показать, что при оценке доли курящих для достижения точности 0,005 с вероятностью 0,95 достаточно провести не 200 000, а всего 40 000 наблюдений.

Упражнение. Решить задачи о проверке гипотезы и об оценке вероятности, используя уточнение теоремы Бернулли из § 5 главы 3 части II. Провести расчеты для рассматривавшихся примеров и сравнить их с полученными. Используя уточнение теоремы Чебышева, попытаться улучшить оценки в примере пункта 3.1.2.

Глава 3

РАЗНЫЕ ЗАДАЧИ

В этой главе рассматриваются некоторые избранные задачи.

§ 1. ЗАДАЧА О РАЗОРЕНИИ ИГРОКА

Игрок, имеющий m рублей, играет n партий в орлянку с партнером, имеющим $c-m$ рублей. Ставка в каждой партии равна одному рублю. Какова вероятность того, что за эти n партий игрок проиграет все свои деньги?

1.1. Математическая формулировка задачи

Рассмотрим произвольные натуральные числа $m, n>0$ и $c>m$. Последовательность n подбрасываний симметричной монеты описывается моделью Бернулли n испытаний с вероятностью успеха $a=1/2$. Для каждого номера $j=1, \ldots, n$ случайная переменная $f_j=2s_j-1$ описывает выигрыш игрока в j-й партии: $f_j(u)=1$, если $u \in У_j$, и $f_j(u)=-1$, если $u \in H_j$. Общий выигрыш игрока, имеющего l рублей $(0<l<c)$ в $j, \ldots, k$-й партиях с партнером, имеющим $c-l$ рублей, описывается случайной переменной g_{kj}^l, определяемой равенствами

$$q_{jj}(u)=f_j(u)$$

$$g_{jk}^l(u)=f_j(u)+g_{j+1,k}^l(u) \qquad (-l<g_{j+1,k}^l(u)<c-l; \quad 1 \leqslant j<k)$$

$$g_{jk}^l(u)=g_{j+1,k}^l(u) \quad \left(g_{j+1,k}^l(u)=-l, \quad c-l; \ 1 \leqslant j<k\right).$$

Замечание. Последние равенства соответствуют формальному продолжению игры, когда один из игроков разорился.

Задача сводится к вычислению вероятности $p(m, n)$ события

$$A=\{u: g_{1n}^m(u)=-m; \quad g_{1k}^m(u) \neq -m, \quad 1 \leqslant k<n\}.$$

2.1. Составление разностного уравнения

Непосредственно подсчитать вероятность $p(m, n)$, по-видимому, трудно. Поэтому с помощью формулы полной вероятности составим разностное уравнение, определяющее вероятность $p(m, n)$.

Рассмотрим события $У_1$ и H_1, описывающие выигрыш и проигрыш игрока в 1-й партии. По формуле полной вероятности

$$(*) \quad P(A)=P(У_1) P_{У_1}(A)+P(H_1) P_{H_1}(A).$$

Имеем: $P(У_1)=a=1/2$ и $P(H_1)=(1-a)=1/2$. Непосредственно из определений вытекает, что при $c-m$, $n>1$

$$У_1A=\{u:f_1(u)=1;\ g_{1n}^m(u)=-m;\ g_{1k}^m(u)\neq -m,\ 1\leqslant k<n\}=$$
$$=\{u:f_1(u)=1;\ g_{2n}^{m+1}(u)=-(m+1);\ g_{2k}^{m+1}(u)\neq-(m+1);$$
$$2\leqslant k<n\}=У_1B.$$

Из независимости испытаний в модели Бернулли следует, что случайные переменные f_1, g_{2n}^{m+1}, независимы и события

$$У_1=\{u:f_1(u)=1\},$$
$$B=\{u:g_{2n}^{m+1}(u)=-(m+1);\quad g_{2k}^{m+1}(u)\neq-(m+1),\quad 2\leqslant k<n\},$$

независимы. Поэтому

$$P_{У_1}(A)=P(У_1A)/P(У_1)=P(У_1B)/P(У_1)=P(У_1)P(B)/P(У_1)=P(B).$$

Из однородности испытаний в модели Бернулли следует, что при $n>1$ вероятность события B равна вероятности события

$$C=\{u:g_{1,n-1}^{m+1}(u)=-(m+1);\ g_{1k}^{m+1}(u)\neq-(m+1),\ 1\leqslant k<n-1\}$$

(проигрыш $m+1$ рублей в 1, ..., $(n-1)$-й партиях). В самом деле, исход $u=u_1u_2\ldots u_n$ принадлежит событию B, если и только если исход $v=u_2\ldots u_nu_1$ принадлежит событию C. Элементарные вероятности исходов u и v равны, следовательно, равны вероятности событий B и C. Таким образом, при $c-m>1$ и $n>1$

$$P_{У_1}(A)=P(B)=p(m+1,\ n-1).$$

Аналогично, при $m>1$ и $n>1$

$$P_{H_1}(A)=p(m-1,\ n-1).$$

Подставляя найденные значения в равенство (*), получаем:

(1) $$p(m,\ n)=(1/2)p(m+1,\ n-1)+(1/2)p(m-1,\ n-1)$$
$$(1<m<c-1,\ 1<n).$$

Кроме того, как нетрудно поверить,

(2) $$p(1,\ n)=(1/2)p(2,\ n-1),\ p(c-1,\ n)=(1/2)p(c-2,\ n-1)$$
$$(1<n),\ p(1,\ 1)=1/2,\ p(m,\ 1)=0\ (1<m<c).$$

Замечание. Равенства (1) и (2) эквивалентны равенствам

$$p(m,\ n)=(1/2)p(m+1,\ n-1)+(1/2)p(m-1,\ n-1)$$
$$(0<m<c,\ 0<n),$$

если

$$p(0,\ 0)=1,\ p(0,\ n)=p(c,\ n)=p(m,\ 0)=0\ (0<m\leqslant c,\ 0<n).$$

Говорят, что числа $p(m, n)$ $(0\leqslant m\leqslant c,\ 0<n)$ образуют решение разностью уравнения (1), удовлетворяющее условиям (2).

3.1. Решение задачи

Используя принцип индукции, нетрудно убедиться в том, что если для произвольных чисел $q(m, n)$ верны равенства

$$q(m, n)=(1/2)q(m+1, n-1)+(1/2)q(m-1, n-1)\ (1<m<c-1,\ 1<n);$$

$$q(1, n)=(1/2)q(2, n-1),\ q(c-1, n)=(1/2)q(c-2, n-1)\ (1<n),$$

$$q(1, 1)=1/2,\ q(m, 1)=0\ (1<m<c),$$

аналогичные равенствам (1) и (2), то

$$q(m, n)=p(m, n)\ (0<m<c,\ 0<n).$$

Замечание. Эквивалентное утверждение выражается фразой: решение уравнения (1), удовлетворяющее условиям (2), единственное.

Оказывается, что единственные числа $p(m, n)$, для которых верны равенства (1) и (2), определяются равенством

$$(3)\quad p(m, n)=\frac{1}{c}\sum_{0<k<c}\sin\frac{km\pi}{c}\cdot\sin\frac{k\pi}{c}\cdot\cos^{n-1}\frac{k\pi}{c}\quad (0<m<c,\ 0<n).$$

Прежде всего заметим, что для каждых числа $x\neq 2\pi l$ $(l=0, \pm1, \pm2, \ldots)$ и номера $n>0$ верно равенство

$$\sum_{1\leqslant k\leqslant n}\cos kx=\frac{1}{2}\left[\sin\left(n+\frac{1}{2}\right)x/\sin\frac{1}{2}x-1\right],$$

которое получается суммированием обеих частей равенства

$$\cos kx\sin\frac{1}{2}x=\frac{1}{2}\left[\sin\left(k+\frac{1}{2}\right)x-\sin\left(k-\frac{1}{2}\right)x\right].$$

Используя указанное равенство для суммы косинусов, получаем:

$$p(1, 1)=\frac{1}{c}\sum_{0<k<c}\sin^2\frac{k\pi}{c}=\frac{1}{2c}\sum_{0<k<c}\left(1-\cos\frac{2\pi}{c}k\right)=$$

$$=\frac{1}{2c}\left\{(c-1)-\frac{1}{2}\left[\sin\left(2\pi-\frac{\pi}{c}\right)\Big/\sin\frac{\pi}{c}-1\right]\right\}=1/2.$$

Если $1<m<c$, то

$$p(m, 1)=\frac{1}{c}\sum_{0<k<c}\sin\frac{km\pi}{c}\sin\frac{k\pi}{c}=\frac{1}{2c}\sum_{0<k<c}\left[\cos\frac{(m-1)\pi}{c}k-\right.$$

$$\left.-\cos\frac{(m+1)\pi}{c}k\right]=\frac{1}{2c}\left\{\sin\left[(m-1)\pi-\frac{(m-1)\pi}{2c}\right]\Big/\sin\frac{(m-1)\pi}{2c}-\right.$$

$$\left.-\sin\left[(m+1)\pi-\frac{(m+1)\pi}{2c}\right]\Big/\sin\frac{(m+1)\pi}{2c}\right\}.$$

Так как числа $m-1$ и $m+1$ оба четные или нечетные, то дроби в фигурных скобках обе равны -1 или $+1$. Поэтому

$$p(m, 1)=0 \quad (1<m<c).$$

Для каждого номера $n>1$ имеем:

$$p(1, n)=\frac{1}{c}\sum_{0<k<c}\sin^2\frac{k\pi}{c}\cos^{n-1}\frac{k\pi}{c}=$$

$$=\frac{1}{2c}\sum_{0<k<c}\sin\frac{2k\pi}{c}\sin\frac{k\pi}{c}\cos^{n-2}\frac{k\pi}{c}=\frac{1}{2}p(2, n-1);$$

$$p(c-1, n)=\frac{1}{c}\sum_{0<k<c}\sin\left(k\pi-\frac{k\pi}{c}\right)\sin\frac{k\pi}{c}\cos^{n-1}\frac{k\pi}{c}=$$

$$=\frac{1}{2c}\sum_{0<k<c}\sin\left(k\pi-\frac{2k\pi}{c}\right)\sin\frac{k\pi}{c}\cos^{n-2}\frac{k\pi}{c}=\frac{1}{2}p(c-2, n-1).$$

Если $1<m<c-1$ и $1<n$, то

$$\frac{1}{2}p(m+1, n-1)+\frac{1}{2}p(m-1, n-1)=\frac{1}{2c}\sum_{0<k<c}\left[\sin\frac{k(m+1)\pi}{c}-\right.$$

$$\left.-\sin\frac{k(m-1)\pi}{c}\right]\sin\frac{k\pi}{c}\cos^{n-2}\frac{k\pi}{c}=\frac{1}{c}\sum_{0<k<c}\sin\frac{km\pi}{c}\sin\frac{k\pi}{c}\cos^{n-1}\frac{k\pi}{c}=$$

$$=p(m, n).$$

Таким образом, из равенств (3) вытекают равенства (1) и (2). Следовательно, равенства (3) определяют искомые вероятности $p(m, n)$.

Замечание. Получение равенства (3) связано с более развитой математической техникой и поэтому не обсуждается. При небольших значениях переменных m и n вероятности $p(m, n)$ можно вычислить, непосредственно используя равенства (1) и (2).

4.1. Случайное блуждание

Вероятностная схема, построенная для задачи о разорении игрока, описывает также *симметричное случайное блуждание частицы по одномерной решетке с поглощающими экранами*. Такая схема иногда используется, например, для описания одномерного *броуновского движения*, при котором частица подвергается ударам со стороны большого числа хаотически двигающихся молекул. Пусть точки $-m, -m+1, \ldots, 0, \ldots, c-m$ оси абсцисс координатной плоскости описывают возможные положения некоторой частицы в моменты $j=1, \ldots, n$. В начальный момент частица находится в точке 0. Если в момент $j-1$ частица находится в точке $x(-m<x<c-m)$, то в момент j она с одинаковой вероятностью переходит в точку $x+1$ или в точку $x-1$. В точках $-m$ и $c-m$ расположены поглощающие экраны: если частица попадает в какую-нибудь из них, то она там и остается.

В этой схеме положение частицы соответствует выигрышу игрока, а поглощение частицы экраном в точке $-m$ — разорению игрока. Равенства (3) определяют вероятность того, что частица в момент n будет поглощена экраном в точке $-m$.

Замечание. Если сформулировать задачу о разорении, не ограничивая заранее число партий и предполагая, что игра ведется до разорения одного из игроков, то при математической постановке задачи для описания такой игры придется рассматривать вероятностное пространство с *бесконечным* множеством исходов.

§ 2. ЗАДАЧА О СПИЧЕЧНЫХ КОРОБКАХ

Известный польский математик С. Банах сформулировал следующую шуточную задачу.

Некто носит с собой две коробки спичек. Время от времени он вынимает спичку из наугад выбранной коробки. Рано или поздно выбранная коробка впервые оказывается пустой. Сколько спичек в этот момент остается в другой коробке?

1.2. Распределение числа остающихся спичек

Обозначим число спичек в каждой из коробок буквой m, а искомое число оставшихся спичек — буквой $x (0 \leqslant x \leqslant m)$. Предположим, что одна из коробок имеет номер 0, а другая — номер 1. Используем модель Бернулли $n = 2m+1$ испытаний с вероятностью успеха $a = 1/2$. Число оставшихся спичек в этой модели описывается случайной переменной f, определяемой так.

В каждой строке u длины $n = 2m+1$, составленной из номеров 0 и 1, число единиц $s(u)$ либо строго больше m, либо строго меньше m. Множество $A = \{u : s(u) > m\}$ строк первого типа описывает появление пустой коробки номер 1, а множество $B = \{u : s(u) < m\}$ строк второго типа — коробки номер 0. В первом случае число остающихся спичек выражается равенством

$$f(u) = m - k(u) \quad (s(u) > m), \tag{1}$$

где $k(u)$ обозначает общее число нулей до $(m+1)$-й единицы в строке u. Во втором случае число остающихся спичек выражается равенством

$$f(u) = m - l(u) \quad (s(u) < m), \tag{2}$$

где $l(u)$ обозначает общее число единиц до $(m+1)$-го нуля в строке u.

Равенства (1) и (2) определяют значение $f(u)$ случайной переменной f для каждого исхода u.

Например, если $m=1$, то

$$f(111)=1-0=1,\ f(110)=1-0=1,\ f(101)=1-1=0,$$
$$f(011)=1-1=0,$$
$$f(000)=1-0=1,\ f(001)=1-0=1,\ f(010)=1-1=0,$$
$$f(100)=1-1=0.$$

Вычислим вероятность $q(x)$ того, что случайная переменная f примет значение x:

$$q(x)=P\{u:f(u)=x\}\quad (0\leqslant x\leqslant m).$$

Это и будет решением задачи.

Для каждой строки u, для которой $s(u)>m$, равенство $f(u)=x$ означает, что:

1) в строке u на $(2m-x+1)$-м месте находится 1;

2) перед ней на $2m-x$ местах произвольно расположены m единиц и $m-x$ нулей;

3) после нее расположена произвольная строка длины x из номеров 0 и 1.

Число таких строк u равно

$$\binom{2m-x}{m}2^x\quad (0\leqslant x\leqslant m).$$

Столько же строк u, для которых $s(u)<m$ и $f(u)=x$. Следовательно, общее число исходов u, для которых $f(u)=x$, равно

$$\binom{2m-x}{m}2^{x+1}\quad (0\leqslant x\leqslant m).$$

Так как исход u имеет вероятность

$$p(u)=2^{-(2m+1)},$$

то

$$q(x)=\binom{2m-x}{m}2^{-(2m-x)}\quad (0\leqslant x\leqslant m).$$

2.2. Среднее число остающихся спичек

Вычислим среднее $E(f)$ случайной переменной f. Непосредственно это сделать, по-видимому, трудно. Поэтому воспользуемся искусственным приемом.

Так как

$$\sum_{0\leqslant x\leqslant m} q(x)=\sum_u p(u)=1,$$

то

$$m - E(f) = \sum_{0 \leqslant x \leqslant m} mq(x) - \sum_{0 \leqslant x \leqslant m} xq(x) = \sum_{0 \leqslant x < m} (m - x)\, q(x) =$$

$$= \sum_{0 \leqslant x < m} (m - x) \binom{2m - x}{m} 2^{-(2m-x)}.$$

Вместе с тем для каждого $x=0, \ldots, m-1$

$$(m - x) \binom{2m - x}{m} = \frac{(m - x)\,(2m - x)!}{m!\,(m - x)!} =$$

$$= \frac{(2m - x)\,(2m - x - 1)!}{m!\,(m - x - 1)!} = (2m - x) \binom{2m - x - 1}{m}.$$

Отсюда вытекает, что

$$m - E(f) = \sum_{0 \leqslant x < m} (2m - x) \binom{2m - x - 1}{m} 2^{-(2m-x)} =$$

$$= \frac{2m+1}{2} \sum_{0 \leqslant x < m} q(x+1) - \frac{1}{2} \sum_{0 \leqslant x \leqslant m} (x+1)\, q(x+1).$$

В то же время

$$\sum_{0 \leqslant x < m} (x+1)\, q(x+1) = \sum_{0 < x \leqslant m} xq(x) = \sum_{0 \leqslant x \leqslant m} xq(x) = E(f),$$

$$\sum_{0 \leqslant x < m} q(x+1) = \sum_{0 < x \leqslant m} q(x) = \sum_{0 \leqslant x \leqslant m} q(x) - q(0) = 1 - q(0).$$

Следовательно,

$$m - E(f) = \frac{2m+1}{2}(1 - q(0)) - \frac{1}{2} E(f).$$

Так как

$$q(0) = \binom{2m}{m} 2^{-2m},$$

то отсюда вытекает, что

$$E(f) = (2m+1) \binom{2m}{m} 2^{-2m} - 1.$$

Это равенство выражает среднее число остающихся спичек.

Замечание. Можно доказать, что

$$E(f) \approx 2\,(m/\pi)^{1/2} - 1.$$

В частности, если $m=12$, то $E(f) \approx 3$; если $m=50$, то $E(f) \approx 7$.

Эти результаты проверяются экспериментально. Например, можно взять две коробки с 12 спичками в каждой и монету. Отметить коробку номер 1 и коробку номер 0. Подбрасывать монету

и, когда она падает гербом вверх, брать спичку из коробки номер 1, а когда цифрой вверх, брать спичку из коробки номер 0. И так до тех пор, пока взятая коробка не окажется пустой. Нужно проделать такой опыт несколько раз, подсчитывая число остающихся спичек. Вследствие закона больших чисел арифметическое среднее этих чисел, как правило, не будет значительно отличаться от числа 3.

Задача. Вычислить дисперсию случайной переменной f.

§ 3. ЗАДАЧА О ДЛИНЕ СЛУЧАЙНОЙ ЦЕПИ

В современной химии для описания молекул некоторых полимеров используется модель плоской случайной цепи, состоящей из линейных звеньев, расположенных под случайными углами друг к другу.

1.3. Определение длины случайной цепи

Представим себе лежащую на столе цепь, составленную из $(n+1)$-го звена. Эти звенья имеют одинаковую длину l_0 и снабжены номерами 0, 1, ..., n. У каждого из звеньев отмечены начало и конец. Поэтому определена величина угла между j-м и $(j-1)$-м звеньями $(1 \leqslant j \leqslant n)$. Эта величина равна либо α, либо $-\alpha$ $(0<\alpha<\pi)$. Углы между звеньями рассматриваемой цепи образуются независимо и одинаково случайно.

Длиной цепи будем считать расстояние между концом последнего n-го звена и началом 0-го звена. Задача заключается в исследовании случайной переменной l, описывающей длину цепи. Проще рассматривать случайную переменную l^2.

Опишем плоскость стола стандартной координатной плоскостью и будем предполагать, что начало координат совпадает с началом 0-го звена, а направление оси x — с направлением этого звена.

Рассмотрим модель Бернулли n испытаний с вероятностью успеха a. Случайная переменная

$$t_k = \left((-1)^{s_1} + \ldots + (-1)^{s_k}\right)$$

описывает величину угла между k-м и 0-м звеньями $(1 \leqslant k \leqslant n)$.

Проекции k-го звена на оси x и y равны соответственно

$$l_0 \cdot \cos t_k, \ l_0 \cdot \sin t_k \ (1 \leqslant k \leqslant n).$$

Следовательно,

$$l^2 = l_0^2 \left[\left(1 + \sum_{1 \leqslant k \leqslant n} \cos t_k\right)^2 + \left(\sum_{1 \leqslant k \leqslant n} \sin t_k\right)^2 \right].$$

Пример. Если $n=3$, $l_0=1$ и $\alpha=2\pi/3$, то значения рассматриваемых случайных переменных содержит таблица

u	11	10	01	00
$t_1(u)$	$-2\pi/3$	$-2\pi/3$	$2\pi/3$	$-2\pi/3$
$t_2(u)$	$-4\pi/3$	0	0	$4\pi/3$
$\cos t_1(u)$	$-1/2$	$-1/2$	$-1/2$	$-1/2$
$\cos t_2(u)$	$-1/2$	1	1	$-1/2$
$\sin t_1(u)$	$-\sqrt{3}/2$	$-\sqrt{3}/2$	$\sqrt{3}/2$	$\sqrt{3}/2$
$\sin t_2(u)$	$\sqrt{3}/2$	0	0	$-\sqrt{3}/2$
$l_2(u)$	0	3	3	0

Квадрат длины цепи l^2 принимает значение 0 с вероятностью $q(0)=a^2+(1-a)^2$ и значение 3 с вероятностью $q(3)=2a(1-a)$.

Замечание. Вычислить вероятности значений случайной переменной l^2 в общем случае, по-видимому, трудно. Пример поясняет рис. 38.

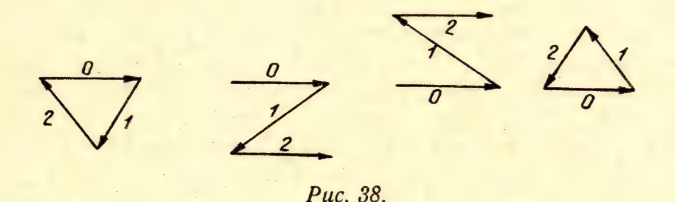


Рис. 38.

2.3. Средний квадрат длины цепи

Вычислим среднее случайной переменной l^2 при $l_0=1$.

1. Докажем сначала, что

$$(1)\quad E[\cos(t_j-t_i)]=\cos^{j-i}\alpha\,(0\leqslant i<n,\ 0<j\leqslant n,\ i<j).$$

Рассмотрим произвольный номер i и проведем доказательство индукцией по j. Если $j=i+1$, то равенство (1) верно, как нетрудно проверить. Если равенство (1) верно для j, то оно верно и для $j+1$. В самом деле,

$$E[\cos(t_{j+1}-t_i)]=E[\cos(t_j-t_i)\cos((-1)^{s_{j+1}}\alpha)]-$$
$$-E[\sin(t_j-t_i)\sin((-1)^{s_{j+1}}\alpha)].$$

Используя четность косинуса, правило сложения для среднего и предположение о верности равенства (1) для j, получаем:

$$E[\cos(t_j-t_i)\cos((-1)^{s_{j+1}}\alpha)]=E[\cos(t_j-t_i)]\cos\alpha=\cos^{(j+1)-i}\alpha.$$

Аналогично, используя нечетность синуса, независимость случайных переменных $\sin(t_j - t_i)$ и $\sin((-1)^{s_{j+1}}\alpha)$, правило умножения для среднего и равенство

$$E[\sin(t_j - t_i)] = 0,$$

которое нетрудно проверить, находим:

$$E[\sin(t_j - t_i)\sin((-1)^{s_{j+1}}\alpha)] =$$
$$= E[\sin(t_j - t_i)]\, E[\sin((-1)^{s_{j+1}}\alpha)] = 0.$$

По принципу индукции из сказанного вытекает, что равенство (1) верно для каждого номера $j > i$.

2. Используя правило сложения для среднего и равенство (1), получаем равенство

$$E(l_j^2) - E(l_{j-1}^2) = 1 + 2\cos\alpha\,\frac{1-\cos^{j-1}\alpha}{1-\cos\alpha}. \tag{2}$$

В самом деле, как нетрудно проверить,

$$E(l_j^2) - E(l_{j-1}^2) = E(l_j^2 - l_{j-1}^2) = E\Big[1 + 2\sum_{0\leqslant i<j}\cos(t_j - t_i)\Big] =$$
$$= 1 + 2\sum_{0\leqslant i<j}\cos^{j-i}\alpha = 1 + 2\cos\alpha\,\frac{1-\cos^{j-1}\alpha}{1-\cos\alpha}.$$

Кроме того, если цепь состоит из $(0+1)$-го звена, то, по предположению, $l_0^2 = 1$ и

$$E(l_0^2) = 1. \tag{2'}$$

3. Складывая равенства (2′) и (2) для $j = 1, \ldots, n$, получаем:

$$E(l_n^2) = 1 + \sum_{1\leqslant j\leqslant n}\frac{1 + 2\cos\alpha\,(1-\cos^{j-1}\alpha)}{1-\cos\alpha}.$$

После некоторых преобразований находим:

$$E(l_n^2) = 1 + n\,\frac{1+\cos\alpha}{1-\cos\alpha} - 2\cos\alpha\,\frac{1-\cos^n\alpha}{(1-\cos\alpha)^2}.$$

Это равенство можно считать решением задачи о длине случайной цепи.

Замечание. Вычислить дисперсию случайной переменной l^2, по-видимому, трудно.

§ 4. ЗАДАЧА О ПЛАНИРОВАНИИ ЭКСПЕРИМЕНТА

Для оценки числа курящих жителей в городе предполагается провести *выборочное анкетирование*. Эксперимент планируется следующим образом: выбираются n из m районов города и в них анкетируются все жители. Обсуждаются схема выбора *без возвращения* и схема выбора *с возвращением*. В схеме выбора без воз-

вращения наугад выбирается один из районов города и туда посылается первый анкетер, затем наугад выбирается один из оставшихся районов и туда посылается второй анкетер и т. д., пока не будет послан последний n-й анкетер. В схеме выбора с возвращением каждый из n анкетеров независимо от других наугад выбирает себе для анкетирования один из m районов города. При этом может случиться, что один и тот же район будет анкетироваться несколько раз, т. е. будут охвачены не n районов, а меньше.

Посылка разных анкетеров в разные районы и охват тем самым точно n районов представляются более разумными. В то же время план, при котором каждый из анкетеров выбирает себе район *наугад*, привлекает своей организационной простотой. Поэтому имеет смысл рассмотреть оба плана подробнее и сравнить оценки среднего значения анкетируемых жителей и оценки случайных отклонений от этого среднего для каждого из планов.

1.4. Схемы выбора с возвращением и без возвращения

Названия схем связаны с тем, что классические формулировки соответствующих задач используют *урновую модель*.

1.1.4. Последовательный выбор с возвращением

В урне находятся m шаров с номерами $1, \ldots, m$. Из нее вынимается наугад выбранный шар и отмечается его номер. После *этого шар возвращается в урну. И так n раз. Какова вероятность вынуть шар с номером i на j-й раз $(m \geqslant 1,\ 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n)$?*

Для математической формулировки этой задачи выберем в качестве множества исходов множество $U = \{1, \ldots, m\}^n$ всех слов $u = u_1 \ldots u_n$ длины n, составленных из номеров $1, \ldots, m$ $(1 \leqslant u_1, \ldots, u_n \leqslant m)$. Из общего правила умножения для числа элементов декартова произведения множеств следует, что число элементов в множестве U равно m^n. Это множество исходов описывает последовательность n испытаний, состоящих каждое в выборе одного из m шаров. Событие $У_{ij} = \{u / u_j = i\}$, составленное из всех слов $u = u_1 \ldots u_n$, у которых на j-м месте находится номер i, соответствует выбору шара с номером i на j-й раз.

В рассматриваемой схеме выбора с возвращением испытания независимы и производятся в одинаковых условиях: состав шаров в урне не меняется. При выборе с возвращением может случиться, что один и тот же шар будет вынут несколько раз и, следовательно, число шаров с различными номерами среди n вынутых будет строго меньше n. Независимость испытаний и одинаковость условий, в которых они проводятся, вместе с предположением о выборе наугад позволяют описать схему выбора с возвращением постоянной элементарной вероятностью p_1 со значениями

(1) $$p_1(u) = 1/m^n \quad (u \in U).$$

Задача сводится к вычислению вероятности $P_1(У_{ij})$ события $У_{ij}$.

Используя *правило умножения*, находим, что число элементов $n(У_{ij})$ в множестве $У_{ij}$ равно m^{n-1}. Поэтому

$$(2) \qquad P_1(У_{ij}) = 1/m \quad (1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n)$$

Замечание. Как и следовало ожидать, эта вероятность не зависит от номера i шара и от номера j испытания.

2.1.4. Последовательный выбор без возвращения

В урне находятся m шаров с номерами $1, \ldots, m$. Из нее вынимается наугад выбранный шар и отмечается его номер. После этого шар не возвращается в урну. И так n раз. Какова вероятность впервые вынуть шар с номером i на j-й раз *$(m \geqslant 1,\ 1 \leqslant n \leqslant m,\quad 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n)$?*

Формулировка отличается от формулировки задачи пункта 1.1.4 добавлением частицы *не* перед словом *возвращается* и условия можно использовать то же множество исходов $U = \{1, \ldots, m\}^n$, выбора без возвращения испытания зависимы и производятся в различных условиях: состав шаров в урне меняется от испытания к испытанию.

Для математического описания схемы выбора без возвращения можно использовать то же множество исходов $U = \{1, \ldots, m\}^n$, что и для схемы выбора с возвращением. Различие этих схем будет описываться различием элементарных вероятностей.

Рассмотрим событие U_0, составленное из всех слов $u = u_1 \ldots u_n$, у которых номера $u_1, \ldots, u_n$ попарно различны. Определим элементарную вероятность p_2, описывающую выбор без возвращения равенствами:

$$(1') \qquad p_2(u) = 1/m \ldots (m-n+1) \quad (u \in U_0),$$
$$p_2(u) = 0 \quad (u \notin U_0).$$

Число элементов $n(U_0)$ в множестве U_0 равно $m \ldots (m-n+1)$. Поэтому сумма всех значений $p_2(u)$ равна 1 и p_2 является элементарной вероятностью для U.

Задача снова сводится к вычислению вероятности $P_2(У_{ij})$ события $У_{ij}$.

Из равенства (1′) вытекает, что

$$P_2(У_{ij}) = P_2(У_{ij} \cap U_0) = n(У_{ij} \cap U_0)/m \ldots (m-n+1).$$

Используя общее правило умножения, нетрудно убедиться в том, что

$$n(У_{ij} \cap U_0) = (m-1) \ldots (m-n+1).$$

Следовательно,

$$(2') \qquad P_2(У_{ij}) = 1/m.$$

Замечание. Таким образом, вероятности вынуть шар с номером i на j-й раз в схемах выбора с возвращением и выбора без

возвращения равны. Эти схемы отличаются условными вероятностями вынуть шар с номером i на j-й раз при условии, что он не был вынут раньше. В схеме выбора с возвращением эта условная вероятность снова равна $1/m$, а в схеме выбора без возвращения — $1/(m-j+1)$.

2.4. Математическая формулировка задачи

Рассмотрим 1-е и 2-е вероятностные пространства $(U, \mathscr{A}, P_1)$ и $(U, \mathscr{A}, P_2)$, определенные в пунктах 1.1.4 и 2.1.4. Опишем 1-й и 2-й планы, использующие выбор с возвращением и выбор без возвращения случайными переменными f_1 и f_2, определенными соответственно для 1-го и 2-го пространства.

Предположим, что число жителей в i-м районе известно и равно $x_i (i=1, \ldots, m)$. Рассмотрим случайную переменную $f_{1j} (j=1, \ldots, n)$ со значениями $f_{1j}(u)=x_i \quad (u \in \mathscr{Y}_{ij})$. Переменная f_{1j} описывает число жителей, анкетированных j-м анкетером при 1-м плане. Сумма

$$f_1 = \sum_{1 \leqslant j \leqslant n} f_{1j}$$

описывает общее число анкетированных жителей при 1-м плане.

Точно так же определяются случайные переменные f_{2j} и f_2 для 2-го плана.

Задача сводится к вычислению средних $E(f_1)$, $E(f_2)$ и дисперсий $D(f_1)$, $D(f_2)$ случайных переменных f_1 и f_2.

Замечание. Верны равенства

$$f_{1j}(u)=f_{2j}(u), \; f_1(u)=f_2(u) \quad (u \in U).$$

Больше того,

$$P_1\{u: f_{1j}(u)=x_i\}=P_2\{u: f_{2j}(u)=x_i\}=1/m \quad (1 \leqslant i \leqslant m, \; 1 \leqslant j \leqslant n),$$

т. е. случайные переменные f_{1j} и f_{2j} одинаково распределены. В то же время случайные переменные f_1 и f_2 имеют различные распределения. Случайные переменные f_{1j} и f_{2j} определены для различных вероятностных пространств, причем переменные f_{1j} независимы, а переменные f_{2j}, вообще говоря, зависимы.

3.4. Решение задачи

Рассмотрим сначала вероятностное пространство $(U, \mathscr{A}, P_1)$ и случайную переменную f_1, затем вероятностное пространство $(U, \mathscr{A}, P_2)$ и случайную переменную f_2.

1.3.4. Среднее и дисперсия для 1-го плана

Используя равенство (2), получаем:

$$E(f_{1j}) = \sum_{1 \leqslant i \leqslant n} x_i P_1(\mathscr{Y}_{ij}) = \frac{1}{m} \sum_{1 \leqslant i \leqslant m} x_i = a \qquad (1 \leqslant j \leqslant n).$$

Поэтому

(3) $$E(f_1) = \sum_{1 \leqslant j < n} E(f_{1j}) = na.$$

Дисперсия для f_{1j} выражается равенствами:

$$D(f_{1j}) = E(f_{1j} - a)^2 = \frac{1}{m} \sum_{1 \leqslant i \leqslant m} (x_i - a)^2 = b^2 \qquad (1 \leqslant j \leqslant n).$$

Нетрудно убедиться в том, что случайные переменные $f_{1j}(1 \leqslant j \leqslant n)$ независимы. Следовательно,

(4) $$D(f_1) = \sum_{1 \leqslant j \leqslant n} D(f_{1j}) = nb^2.$$

2.3.4. Среднее и дисперсия для 2-го плана

Используя равенство (2'), получаем:

$$E(f_{2j}) = \sum_{1 \leqslant i \leqslant n} x_i P_2(\mathcal{Y}_{ij}) = \frac{1}{m} \sum_{1 \leqslant i \leqslant m} x_i = a \qquad (1 \leqslant j \leqslant n).$$

Поэтому

(3') $$E(f_2) = \sum_{1 \leqslant j \leqslant n} E(f_{2j}) = na.$$

Вычисление дисперсии для f_2 связано с более длинными выкладками. Дисперсия для f_{2j} так же, как и для f_{1j}, выражается равенством

$$D(f_2) = b^2 \quad (1 \leqslant j \leqslant n).$$

Вычислим средние произведений $(f_{2j} - a)\ (f_{2l} - a)$ при $1 \leqslant j < l \leqslant n$. Рассмотрим произвольные различные номера i, k между 1, m и множество $\{u : f_{2j}(u) = x_i,\ f_{2l}(u) = x_k\} \cap U_0$. Используя правило умножения, нетрудно проверить, что число элементов этого множества равно $(m-2) \ldots (m-n+1)$. Следовательно,

$$P_{j,\,l}(i, k) = P_2(\{u : f_{2j}(u) = x_i,\ f_{2l}(u) = x_k\}) = 1/m(m-1) \quad (i \neq k).$$

Если $i = k$, то соответствующая вероятность равна 0. Следовательно,

$$E[(f_{2j} - a)(f_{2l} - a)] = \sum_{i,k} (x_i - a)(x_k - a) P_{j,l}(i, k) =$$

$$= \frac{1}{m(m-1)} \sum_{i \neq k} (x_i - a)(x_k - a) = \frac{1}{m(m-1)} \left\{ \left[\sum_{1 \leqslant i \leqslant m} (x_i - a) \right]^2 - \right.$$

$$\left. - \sum_{1 \leqslant i \leqslant m} (x_i - a)^2 \right\} = -\frac{1}{m-1} b^2.$$

Наконец, используя определение дисперсии и полученные равенства, находим:

$$D(f_2) = E(f_2 - na)^2 = E\left[\sum_{1\leqslant j\leqslant n} (f_{2j} - a)\right]^2 = E\left[\sum_{1\leqslant j\leqslant n} (f_{2j} - a)^2 + \right.$$

$$\left. + 2 \sum_{1\leqslant j<l\leqslant n} (f_{2j} - a)(f_{2l} - a)\right] = \sum_{1\leqslant j\leqslant n} D(f_{2j}) +$$

$$+ 2 \sum_{1\leqslant j<l\leqslant n} E[(f_{2j} - a)(f_{2l} - a)] = nb^2 - 2\binom{n}{2}\frac{1}{m-1} b^2.$$

Таким образом,

$$D(f_2) = \frac{n(m-n)}{m-1} b^2. \tag{4'}$$

Сравнивая равенства (3) и (3′), (4) и (4′), видим, что в среднем планы 1-й и 2-й равноценны: при каждом из них можно ожидать, что число анкетированных жителей равно na, где a — среднее число жителей в районе. Однако при плане 1 с хаотическим выбором районов можно ожидать бо́льших отклонений от числа na, чем при плане 2 с более упорядоченным выбором районов.

Замечание. Из полученных равенств вытекает, что если $x_i \neq a$ для некоторых номеров i и $b \neq 0$, то случайные переменные f_{2j} и f_{2l} зависимы и коэффициент корреляции для них выражается равенством

$$K(f_{2j}, f_{2l}) = \frac{1}{m-1} \qquad (j \neq l).$$

Равенство $b^2 = 0$ означает, что в каждом районе города ровно a жителей.

§ 5. ЗАДАЧА ОБ АНАЛИЗЕ КРОВИ

Сравним следующие два метода анализа крови для группы $n = k \cdot l$ человек. У каждого из этих людей результат анализа может быть либо *отрицательным*, либо *положительным*.

При первом методе кровь каждого человека анализируется отдельно. Ясно, что для этого требуется ровно n анализов.

При втором методе группу разбивают на l подгрупп по k человек. Часть взятой для анализа крови каждого из этих k человек смешивается, полученная смесь анализируется. Если результат анализа смеси отрицателен, то этого одного анализа достаточно: можно утверждать, что результаты для всех k человек отрицательны. Если же результат анализа смеси положителен, то проводится дополнительный анализ для каждого из этих людей. В этом случае для группы из k человек проводится $1+k$ анализов.

Требуется оценить среднее количество анализов при втором методе.

1.5. Математическая формулировка задачи

Будем предполагать, что рассматриваемая группа из n людей описывается моделью Бернулли n испытаний с вероятностью успеха a, считая успехом отрицательный результат анализа. Произведем нумерацию людей в группе и подгруппах, а также нумерацию самих подгрупп. Обозначим $n(i, j)$ номер в группе i-го человека j-й подгруппы.

Случайная переменная f_{ij} со значениями

$$f_{ij}(u) = 1 - s_{n(i,\ j)}(u) \quad (1 \leqslant i \leqslant k,\ 1 \leqslant j \leqslant l)$$

описывает результат анализа i-го человека j-й подгруппы: она равна 0 в случае отрицательного и 1 в случае положительного результата анализа для этого человека.

Случайная переменная

$$f_j = \sum_{1 \leqslant i \leqslant j} f_{ij} \qquad (1 \leqslant j \leqslant l)$$

описывает результат анализа смеси для j-й подгруппы: эта переменная равна 0, если результат анализа отрицателен, и не равна 0, если он положителен.

Для каждого вещественного числа x положим:

$$\operatorname{sgn} x = -1 \ (x < 0),\ \operatorname{sgn} 0 = 0,\ \operatorname{sgn} x = 1 \ (x > 0).$$

Случайная переменная g_j со значениями

$$g_j(u) = 1 + k \cdot \operatorname{sgn} f_j(u) f_{ij}(u) \quad (1 \leqslant j \leqslant l)$$

описывает число анализов для j-й подгруппы: она равна 1, если результат анализа смеси отрицателен, и равна $1+k$, если положителен. Наконец, случайная переменная

$$g = \sum_{1 \leqslant j \leqslant l} g_j$$

описывает общее число анализов для второго метода. Задача сводится к вычислению среднего $E(g)$ случайной переменной g.

2.5. Решение задачи

Заметим, что

$$Eg = E\left(\sum_{1 \leqslant j \leqslant l} g_j\right) = \sum_{1 \leqslant j \leqslant l} E(g_j).$$

Кроме того,

$$E(g_j) = E(1 + k \cdot \operatorname{sgn} f_j) = 1 + kE(\operatorname{sgn} f_j) \quad (1 \leqslant j \leqslant l).$$

Таким образом, дело сводится к вычислению среднего $E(h_j)$ случайных переменных h_j со значениями

$$h_j(u) = \operatorname{sgn} f_j(u) \quad (1 \leqslant j \leqslant l).$$

Случайная переменная h_j принимает значения 0 и 1. Следовательно,

$$E(h_j)=P\{u:h_j(u)=1\}=1-P\{u:h_j(u)=0\} \quad (1\leqslant j\leqslant l).$$

Равенство $h_j(u)=0$ эквивалентно равенству $f_j(u)=0$, которое эквивалентно равенствам

$$f_{1j}(u)=\ldots=f_{kj}(u)=0 \quad (1\leqslant j\leqslant l).$$

Эти равенства, в свою очередь, эквивалентны равенствам

$$s_{n(1,j)}(u)=\ldots s_{n(k,j)}(u)=1 \quad (1\leqslant j\leqslant l).$$

Так как

$$P\{u:s_{n(1,j)}(u)=\ldots=s_{n(k,j)}(u)=1\}=a^k \quad (1\leqslant j\leqslant l),$$

то, следовательно,

$$E(h_j)=1-a^k \quad (1\leqslant j\leqslant l)$$

и

$$E(g)=l[1+k(1-a^k)].$$

Используя равенство $l=n/k$, получаем

$$E(g)=n[1-a^k+1/k].$$

Замечание. Если при определении случайной переменной, описывающей число анализов, не вдаваться в подробности, то решение задачи выглядит очень просто. В самом деле, если анализы крови отдельных людей независимы, то для каждой подгруппы вероятность отрицательного результата анализа смеси равна a^k, а положительного —$1-a^k$. Число анализов соответственно равно 1 и $1+k$. Следовательно, среднее числа анализов для каждой группы равно

$$1\cdot a^k+(1+k)(1-a^k)=1+k(1-a^k).$$

Так как подгрупп l, то среднее общего числа анализов равно

$$l[1+k(1-a^k)]=n(1-a^k+1/k).$$

Если вероятность a отрицательного результата анализа достаточно велика ($a^k>1/k$), то $E(g)<n$ и второй метод в среднем более экономичен. Например, если $a=0{,}8$ и $k=2$, то

$$E(g)=n(1-0{,}14).$$

В этом случае можно ожидать, что второй метод дает по сравнению с первым среднюю экономию около 14%. Для больших групп это может иметь существенное значение.

Замечание. Во время второй мировой войны описанный метод применялся в армейских условиях и давал экономию в числе анализов до 80%.

§ 6. ЗАДАЧА О НАИБОЛЬШЕЙ ДИСПЕРСИИ

Модель Бернулли описывает последовательность независимых и одинаково распределенных описаний с двумя случайными исходами каждое. Естественно рассмотреть аналогичную модель, описывающую произвольную последовательность n независимых испытаний с двумя случайными исходами каждое. Для такой модели так же, как и для модели Бернулли, определено *общее число успехов*. Возникает задача о том, для какой из этих моделей с *данной средней вероятностью успеха* общее число успехов имеет наибольшую дисперсию?

1.6. Модель Бернулли n испытаний с вероятностями успеха $a_1, \ldots, a_n$

В качестве *множества исходов* для этой модели выберем то же самое множество, что и для обычной модели Бернулли: множество U всех слов $u=u_1 \ldots u_n$ длины n, составленных из чисел 0 и 1. Для новой модели определены те же события и переменные, что и для обычной модели Бернулли. В частности, определены события $У_j$ и H_j — *успех* при j-м испытании и *неудача* при j-м испытании, а также связанные с ними переменные s_j и $s=\Sigma s_j$ — число успехов при j-м испытании и общее число успехов.

Рассмотрим числа $a_1, \ldots, a_n$, удовлетворяющие условию $0 \leqslant a_j \leqslant 1$ $(n \geqslant 1,\ j=1, \ldots, n)$. Определим *элементарную вероятность* p равенством

$$p(u) = \Pi a_j^{s_j(u)} (1-a_j)^{1-s_j(u)} \quad (u \in U).$$

Используя принцип индукции, нетрудно убедиться в том, что p является элементарной вероятностью для U. В частном случае, когда $a_1 = \ldots = a_n = a$, эта элементарная вероятность равна элементарной вероятности для обычной модели Бернулли n испытаний с вероятностью успеха a.

Множество исходов U и элементарная вероятность p, выбранные указанным образом, определяют вероятностную модель $(U, \mathscr{A}; p, P)$. Условимся называть эту модель моделью Бернулли n испытаний с вероятностями успеха $a_1, \ldots, a_n$. Обычная модель Бернулли n испытаний с вероятностью успеха a является частным случаем этой новой модели при $a_1 = a_2 = \ldots = a_n = a$.

Как нетрудно проверить, в новой модели

$$P\{u : s_j(u) = 1\} = P(У_j) = a_j,$$
$$P\{u : s_j(u) = 0\} = P(H_j) = 1 - a_j$$

и, следовательно,

$$E(s_j) = a_j, \ D(s_j) = a_j(1-a_j).$$

Нетрудно убедиться также в том, что случайные переменные s_j по-прежнему независимы. Поэтому

$$E(s) = \Sigma a_j, \ D(s) = \Sigma a_j(1-a_j).$$

Если $a_1 = \ldots = a_n = a$, то эти равенства эквивалентны соответствующим равенствам для обычной модели Бернулли n испытаний с вероятностью успеха a.

2.6. Решение задачи о наибольшей дисперсии

Рассмотрим модель Бернулли n испытаний с вероятностями успеха $a_1, \ldots, a_n$. Назовем число $a = (1/n)\Sigma a_j$ *средней вероятностью* успеха для этой модели. В частности, если $a_1 = \ldots = a_n = a$, то средняя вероятность успеха совпадает с обычной. Как уже отмечалось,

$$E(s) = na, \quad D(s) = na - \sum a_j^2.$$

Рассмотрим произвольные номер $n \geqslant 1$ и число a. Задача состоит в том, чтобы среди всех моделей Бернулли n испытаний с вероятностями успеха $a_1, \ldots, a_n$, у которых средняя вероятность равна a, найти модели с наибольшей дисперсией $D(s) = na - \sum a_j^2$ общего числа успехов s.

Заметим, что

$$\sum (a_j - a)^2 = \sum \left(a_j^2 - 2aa_j + a^2\right) =$$
$$= \sum a_j^2 - 2a \sum a_j + na^2 = \sum a_j^2 - na^2.$$

Следовательно,

$$\sum a_j^2 = na^2 + \sum (a_j - a)^2$$

и сумма $\sum a_j^2$ является наименьшей, а дисперсия $D(s)$ — наибольшей, если и только если $a_1 = a_2 = \ldots = a_n = a$.

Замечание. Таким образом, модель Бернулли n испытаний с вероятностями успеха $a_1 = \ldots = a_n = a$ является единственной моделью среди всех моделей Бернулли n испытаний с вероятностями успеха $a_1, \ldots, a_n$, имеющих среднюю вероятность успеха, равную a, для которой дисперсия $D(s)$ общего числа успехов s наибольшая. Этот результат представляется довольно неожиданным и приводит к парадоксальным, на первый взгляд, выводам.

Рассмотрим, например, систему из независимых элементов с двумя состояниями, характеризующуюся средней вероятностью одного из этих состояний. Оказывается, что если система составлена из однородных элементов, то можно ожидать наибольших случайных отклонений от среднего режима. Обычно в этом случае ожидают наименьших отклонений.

§ 7. ЗАДАЧА О СЛУЧАЙНОМ БЛУЖДАНИИ

Пусть имеются белая и черная урны с белыми и черными шарами. Доля белых шаров в белой урне $q_{00} = 1 - b$, а доля черных шаров в белой урне $q_{10} = b\,(0 < b < 1)$. Аналогично, доля бе-

урне $q_{11}=1-a$ $(0<a<1)$. Кроме того, пусть имеется еще одна лых шаров в черной урне $q_{01}=a$, а доля черных шаров в черной урна — красная, тоже с белыми и черными шарами. Доля белых шаров в красной урне $p_{00}=c$, а доля черных шаров в красной урне $p_{10}=1-c$ $(0<c<1)$.

Рассмотрим следующую последовательность $1+n$ испытаний $(n\geqslant1)$; начальное 0-е испытание состоит в выборе наугад шара из красной урны; k-е испытание состоит в выборе наугад шара из урны цвета $(k-1)$-го вынутого шара, регистрации цвета этого k-го шара и возвращении его в ту же урну $(k=1, \ldots, n)$. В частности, если 0-й шар, вынутый из красной урны, оказался белым, то 1-й шар вынимается наугад из белой урны и после регистрации цвета возвращается в нее; если 0-й шар оказался черным, то 1-й шар вынимается наугад из черной урны и после регистрации цвета возвращается в нее.

Возникает задача: *какова вероятность того, что k-й вынутый шар будет белым*?

1.7. Математическая формулировка задачи

Ясно, что в модели, описывающей рассматриваемую последовательность испытаний, вероятность вынуть белый шар при начальном 0-м испытании должна быть равна p_{00}, а черный — p_{10}. Точно так же ясно, что условная вероятность вынуть белый шар при k-м испытании, если при $(k-1)$-м вынут белый шар, должна быть равна q_{00}, а черный — q_{10}. Аналогично, условная вероятность того, что k-й шар белый, если $(k-1)$-й — черный, должна равняться q_{01}, а того, что черный — q_{11} $(k=1, \ldots, n)$. Характерной особенностью рассматриваемой последовательности является *цепная зависимость* испытаний: вероятности исходов k-го испытания определяются исходом $(k-1)$-го.

Определим вероятностную модель, описывающую рассматриваемую последовательность испытаний, следующим образом. В качестве *множества исходов* возьмем множество U всех слов $u=u_0u_1\ldots u_n$ длины $1+n$, составленных из номеров 0 и 1. *Элементарную вероятность* p определим равенством:

$$(1)\quad p(u)=q_{i_n i_{n-1}}\ldots q_{i_1 i_0}p_{i_0 0}\ (n\geqslant1,\ 0\leqslant i_0, i_1, \ldots, i_n\leqslant1).$$

Будем предполагать, что

$$(1')\quad p_{00},\ p_{10}>0,\ p_{00}+p_{10}=1;\ q_{00},\ q_{10}>0,\ q_{00}+q_{10}=1;\ q_{01},\ q_{11}>0,\ q_{01}+q_{11}=1.$$

Если $n=1$, то $U=\{00, 01, 10, 11\}$ и

$$p(00)=q_{00}p_{00},\ p(01)=q_{10}p_{00},\ p(10)=q_{01}p_{10},\ p(11)=q_{11}p_{10}.$$

Для того, чтобы проверить, что равенство (1) действительно определяет элементарную вероятность, достаточно заметить, что из

условий (1′) вытекают условия $p(u)>0$, $\Sigma p(u)=1$. Последнее равенство легко доказать, используя принцип индукции.

Убедимся в том, что определенная такими множеством исходов U и элементарной вероятностью p вероятностная модель обладает нужными для описания рассматриваемой последовательности свойствами. Условимся белый цвет обозначать номером 0, а черный — номером 1. В частности, событие *белый шар при k-м испытании* будем обозначать равенством $i_k=0$, а событие *черный шар при k-м испытании* — равенством $i_k=1$. При таком соглашении символ $P(i_0=0)$, например, обозначает вероятность того, что при 0-м испытании вынут белый шар, а символ $P(i_0=1)$,— черный. В рассматриваемой модели испытания и все связанные с ними понятия определяются так же, как и для модели Бернулли.

Рассмотрим сначала частный случай. При $n=1$ имеем:

$$P(i_0=0)=P(\{00,\ 01\})=p(00)+p(01)=q_{00}p_{00}+q_{10}p_{00}=$$
$$=(q_{00}+q_{10})p_{00}=p_{00},$$

$$P(i_0=1)=P(\{10,\ 11\})=p(10)+p(11)=q_{01}p_{10}+q_{11}p_{10}=$$
$$=(q_{01}+q_{11})p_{10}=p_{10}.$$

Таким образом, вероятность вынуть шар i-го цвета при начальном 0-м испытании действительно равна p_{i0} $(i=0,\ 1)$. Кроме того, учитывая условие $p_{10},\ p_{00}>0$, получаем:

$$P_{i_0=0}(i_1=0)=\frac{P(i_0=0,\ i_1=0)}{P(i_0=0)}=\frac{p(00)}{p_{00}}=\frac{q_{00}p_{00}}{p_{00}}=q_{00},$$

$$P_{i_0=0}(i_1=1)=\frac{P(i_0=0,\ i_1=1)}{P(i_0=0)}=\frac{p(01)}{p_{00}}=\frac{q_{10}p_{00}}{p_{00}}=q_{10},$$

$$P_{i_0=1}(i_1=0)=\frac{P(i_0=1,\ i_1=0)}{P(i_0=1)}=\frac{p(10)}{p_{10}}=\frac{q_{01}p_{10}}{p_{10}}=q_{01},$$

$$P_{i_0=1}(i_1=1)=\frac{P(i_0=1,\ i_1=1)}{P(i_0=1)}=\frac{p(11)}{p_{10}}=\frac{q_{11}p_{10}}{p_{10}}=q_{11}.$$

Таким образом, условная вероятность вынуть шар i-го цвета при k-м испытании, если при $(k-1)$-м вынут шар j-го цвета, действительно равна q_{ij} $(0\leqslant i,\ j\leqslant 1,\ 1\leqslant k\leqslant n)$.

В общем случае, аналогичные выкладки с учетом условий $p_{j0},\ q_{ij}>0$ приводят к тем же результатам:

$$P(i_0=j)=\sum q_{i_n,i_{n-1}}\ldots q_{i_1,j}p_{j0}=p_{j0}\sum q_{i_n,i_{n-1}}\ldots q_{i_1,j}=p_{j0},$$

$$P_{i_{k-1}=j}(i_k=i)=\frac{P(i_{k-1}=j,\ i_k=i)}{P(i_{k-1}=j)}=$$

$$=\frac{q_{ij}\sum q_{i_n,i_{n-1}}\ldots q_{i_{k+1},i}q_{j,i_{k-2}}\ldots q_{i_1,i_0}p_{i_00}}{\sum q_{i_n,i_{n-1}}\ldots q_{i_k,j}q_{j,i_{k-2}}\ldots q_{i_1,i_0}p_{i_00}}=q_{ij}.$$

Суммирование ведется по всем индексам $0 \leqslant i_0, i_1, \ldots, i_n \leqslant 1$, кроме $i_0 = j$ для первой строки, $i_{k-1} = j$, $i_k = i$ и $i_{k-1} = j$ — соответственно для числителя и знаменателя третьей. Равенство

$$\sum_{0 \leqslant i_k, \ldots, i_n \leqslant 1} q_{i_n, i_{n-1}} \ldots q_{i_k, j} = 1 \qquad (j = 0, 1;\ k = 1, \ldots, n),$$

которое используется в обеих строках, легко доказывается по индукции.

Мы доказали, что в рассматриваемой вероятностной модели вероятность j-го исхода при начальном 0-м испытании равна p_{j0}. Кроме того, было доказано, что условная вероятность i-го исхода k-го испытания при j-м исходе $(k-1)$-го испытания равна q_{ij} $(0 \leqslant i, j \leqslant 1, 1 \leqslant k \leqslant n)$ (подчеркнем, что эти вероятности не зависят от номера испытания). Таким образом, рассматриваемая вероятностная модель обладает нужными для описания рассматриваемой последовательности испытаний свойствами.

Говорят, что эта модель описывает *марковскую цепь* длины n. Номера 0 и 1 называют ее *состояниями*. Числа $p_{i_0 0}$ — *начальными*, а q_{ij} — *переходными вероятностями* цепи. Вместо испытаний говорят о *моментах времени*.

Задача состоит в вычислении вероятностей

$$p_{0k} = P(i_k = 0),\ p_{1k} = P(i_k = 1)$$

состояний цепи в k-й момент $(k = 1, \ldots, n)$.

2.7. Матричная запись

Рассмотрим *матрицы*

$$\bar{p}_0 = \begin{pmatrix} p_{00} \\ p_{10} \end{pmatrix} = \begin{pmatrix} c \\ 1-c \end{pmatrix}, \quad Q = \begin{pmatrix} q_{00} q_{01} \\ q_{10} q_{11} \end{pmatrix} = \begin{pmatrix} 1-b & a \\ b & 1-a \end{pmatrix}.$$

Матрица $\bar{p}_0$ называется вектором *начальных* вероятностей, а матрица Q — матрицей *переходных* вероятностей. Рассмотрим также вектор вероятностей в k-й момент:

$$\bar{p}_k = \begin{pmatrix} p_{0k} \\ p_{1k} \end{pmatrix} \quad (k = 1, \ldots, n).$$

Определим умножение матриц правилом *строка на столбец*. Например:

$$Q\bar{p}_{k-1} = \begin{pmatrix} q_{00} q_{01} \\ q_{10} q_{11} \end{pmatrix} \begin{pmatrix} p_{0,k-1} \\ p_{1,k-1} \end{pmatrix} = \begin{pmatrix} q_{00} p_{0,k-1} + q_{01} p_{1,k-1} \\ q_{10} p_{0,k-1} + q_{11} p_{1,k-1} \end{pmatrix} \quad (k = 1, \ldots, n),$$

$$Q^2 = QQ = \begin{pmatrix} q_{00} q_{01} \\ q_{10} q_{11} \end{pmatrix} \begin{pmatrix} q_{00} q_{01} \\ q_{10} q_{11} \end{pmatrix} = \begin{pmatrix} q_{00} q_{00} + q_{01} q_{10} q_{00} q_{01} + q_{01} q_{11} \\ q_{10} q_{00} + q_{11} q_{10} q_{10} q_{01} + q_{11} q_{11} \end{pmatrix}.$$

Используя условия q_{i0}, $q_{i1}>0$, $p_{0,k-1}=P(i_{k-1}=0)>0$, $p_{1,k-1}=P(i_{k-1}=1)>0$ и формулу полной вероятности, получаем:

$$q_{i0}p_{0,k-1}+q_{i1}p_{1,k-1}=P_{i_{k-1}=0}(i_k=i)P(i_{k-1}=0)+$$
$$+P_{i_{k-1}=1}(i_k=i)P(i_{k-1}=1)=p_{ik}.$$

Понимая под равенством матриц равенство их соответствующих элементов и сравнивая полученные результаты, видим, что

$$\overline{p_k}=Q\overline{p_{k-1}}\quad (k=1,\ldots,n).$$

Используя это равенство последовательно, убеждаемся в том, что

(2) $$\overline{p_k}=Q^k\overline{p_0}\quad (k=1,\ldots,n).$$

Это равенство связывает вероятности состояний в k-й момент с начальными и переходными вероятностями, и его можно считать решением поставленной задачи. В рассматриваемом простом случае удается пойти дальше.

3.7. Решение задачи

Как нетрудно проверить,

$$Q=\begin{pmatrix}1-b & a\\ b & 1-a\end{pmatrix}=\frac{1}{a+b}\begin{pmatrix}a & a\\ b & b\end{pmatrix}+\frac{1-a-b}{a+b}\begin{pmatrix}b & -a\\ -b & a\end{pmatrix}.$$

(Умножение матрицы на число означает умножение каждого ее элемента на это число.)

Например:

$$\frac{1}{a+b}a+\frac{1-a-b}{a+b}b=\frac{a+(1-a-b)b}{a+b}=\frac{(a+b)(1-b)}{a+b}=1-b,$$
$$\frac{1}{a+b}b+\frac{1-a-b}{a+b}(-b)=\frac{b-(1-a-b)b}{a+b}=\frac{(a+b)b}{a+b}.$$

Заметим, что

$$\begin{pmatrix}a & a\\ b & b\end{pmatrix}\begin{pmatrix}b & -a\\ -b & a\end{pmatrix}=\begin{pmatrix}ab-ab & -aa+aa\\ bb-bb & -ab+ab\end{pmatrix}=\begin{pmatrix}00\\ 00\end{pmatrix}=0.$$

Кроме того,

$$\begin{pmatrix}a & a\\ b & b\end{pmatrix}^2=\begin{pmatrix}a & a\\ b & b\end{pmatrix}\begin{pmatrix}a & a\\ b & b\end{pmatrix}=\begin{pmatrix}a^2+ab & a^2+ab\\ ab+b^2 & ab+b^2\end{pmatrix}=(a+b)\begin{pmatrix}b & -a\\ -b & a\end{pmatrix}.$$

Используя эти равенства и последовательно вычисляя $Q^2,\ldots,Q^k$, получаем:

(3) $$Q^k=\frac{1}{a+b}\begin{pmatrix}a & a\\ b & b\end{pmatrix}+\frac{(1-a-b)^k}{a+b}\begin{pmatrix}b & -a\\ -b & a\end{pmatrix}\quad (k=1,\ldots,n).$$

Заметим, что

$$\begin{pmatrix} a & a \\ b & b \end{pmatrix}\begin{pmatrix} c \\ 1-c \end{pmatrix}=\begin{pmatrix} ac+a(1-c) \\ bc+b(1-c) \end{pmatrix}=\begin{pmatrix} a \\ b \end{pmatrix},$$

$$\begin{pmatrix} b & -a \\ -b & a \end{pmatrix}\begin{pmatrix} c \\ 1-c \end{pmatrix}=\begin{pmatrix} bc-a(1-c) \\ -bc+a(1-c) \end{pmatrix}=[(a+b)c-a]\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Отсюда и из равенств (2) и (3) следует равенство

$$(4)\qquad \bar{p}_k=\frac{1}{a+b}\begin{pmatrix} a \\ b \end{pmatrix}+(1-a-b)^k\left(c-\frac{a}{a+b}\right)\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Это равенство эквивалентно равенствам

$$p_{0k}=\frac{a}{a+b}+(1-a-b)^k\left(c-\frac{a}{a+b}\right),$$

$$(4')\qquad p_{1k}=\frac{b}{a+b}-(1-a-b)^k\left(c-\frac{a}{a+b}\right).$$

Равенства (4′) и являются решением задачи: первое из них выражает вероятность состояния 0, а второе — состояния 1 в k-й момент ($k=1,\ldots,n$).

Так как $0<a,\ b<1$, то $|1-a-b|<1$. Поэтому при достаточно больших номерах k члены с $(1-a-b)^k$ произвольно малы и ими можно пренебрегать. Таким образом,

$$p_{0k}\approx\frac{a}{a+b}=\frac{q_{01}}{q_{01}+q_{10}},\quad p_{1k}\approx\frac{b}{a+b}=\frac{q_{10}}{q_{01}+q_{10}}.$$

Полученные результаты показывают, что с течением времени влияние начального состояния ослабевает и вероятности состояний начинают определяться главным образом переходными вероятностями цепи. Для урновой модели это означает, что вероятность вынуть белый шар при испытании с достаточно большим номером практически определяется составом шаров в белой и черной урнах и мало зависит от состава шаров в красной урне.

4.7. Модель Бернулли

Если $a+b=1$, то из равенств (4′) следует, что

$$p_{0k}=a,\ p_{1k}=1-a\ (k=1,\ldots,n).$$

В этом случае матрица переходных вероятностей

$$(*)\qquad Q=\begin{pmatrix} a & a \\ 1-a & 1-a \end{pmatrix}$$

имеет одинаковые столбцы. Для урновой модели это означает, что состав шаров в белой и черной урнах одинаков и можно обойтись одной из них.

Обозначим $s(u)$ число номеров 1 в слове $u=i_1, \ldots, i_n$. В рассматриваемом случае верны равенства:

$$P\{0i_1, \ldots, i_n, 1i_1, \ldots, i_n\} = ca^{s(u)}(1-a)^{n-s(u)} + (1-c)a^{s(u)}(1-a)^{n-s(u)} = $$
$$= a^{s(u)}(1-a)^{n-s(u)}.$$

Следовательно, дело сводится к модели Бернулли n испытаний с вероятностью успеха a.

Из равенств (4′) следует также, что при $n>1$ и $0<a<1$ такая модель Бернулли определяет единственную марковскую цепь рассматриваемого вида, для которой испытания независимы. В самом деле, если k-е и $(k-1)$-е испытания независимы, то

$$a = q_{01} = P_{i_{k-1}=1}(i_k = 0) = P(i_k = 0) =$$
$$= p_{0k} = \frac{a}{a+b} + (1-a-b)^k\left(c - \frac{a}{a+b}\right) \quad (k = 1, 2, \ldots, n).$$

Эти равенства возможны, только если правая часть не зависит от k, т. е. только если второе слагаемое в ней равно 0 (равенство $1-a-b=1$ эквивалентно равенствам $a=b=0$ и поэтому исключено). В этом случае $a=a/(a+b)$ и, следовательно, $a+b=1$.

Таким образом, схема Бернулли n испытаний с вероятностью успеха a при $n>1$ и $0<a<1$ эквивалентна марковской цепи длины n с двумя состояниями, произвольными начальными вероятностями и матрицей переходных вероятностей, определяемой равенством (*).

5.7. Случайное блуждание

Наглядной иллюстрацией марковской цепи с двумя состояниями является следующая физическая схема. В начальный момент 0 частица с вероятностью $p_{10}=c$ находится в точке 1, а с вероятностью $p_{20}=1-c$ — в точке 2 $(0<c<1)$. Если в момент $k-1$ частица находится в точке 1, то в момент k она с вероятностью $q_{11}=1-b$ остается в точке 1, а с вероятностью $q_{21}=b$ перескакивает в точку 2 $(0<b<1)$. Аналогично, если в момент $k-1$ частица находится в точке 2, то в момент k она с вероятностью $q_{12}=a$ перескакивает в точку 1, а с вероятностью $q_{22}=1-a$ остается в точке 1 $(0<a<1)$. Этот процесс продолжается до момента n $(n>1, k=1, \ldots, n)$.

Какова вероятность того, что в момент k частица находится в точке 1?

Описанную схему случайного блуждания частицы по двум точкам можно представить следующей диаграммой (рис. 39).

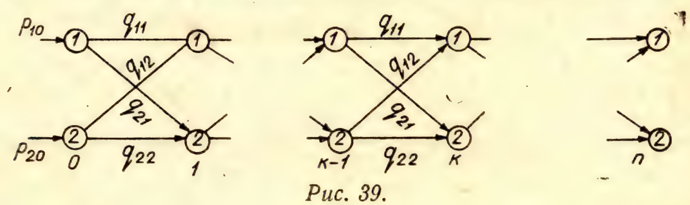


Рис. 39.

На этой диаграмме пути, составленные из стрелок, изображают возможные пути частицы во времени и пространстве. Перемножая вероятности, которыми взвешены соответствующие стрелки, получаем элементарную вероятность каждого данного пути.

Ответ на поставленный вопрос дается первым из равенств (4′):

$$p_{1k} = \frac{a}{a+b} + (1-a-b)^k \left(c - \frac{a}{a+b}\right) \quad (k=1, \ldots, n).$$

Например, если $a=b=c=1/4$, то

$$p_{1k} = \frac{1}{2} - \frac{1}{2^{k+2}} \qquad (k=1, \ldots, n).$$

6.7. Обучение

Простейшую модель обучения крысы определенной реакции на раздражение можно описать марковской цепью с двумя состояниями. На начальное 0-е раздражение крыса с вероятностью $p_{10}=c$ реагирует правильно, а с вероятностью $p_{20}=1-c$ — неправильно $(0<c<1)$. Если на $(k-1)$-е раздражение крыса реагирует правильно, то после обучения на k-е раздражение она с вероятностью $q_{11}=1-b$ реагирует правильно, а с вероятностью $q_{21}=b$ — неправильно $(0<b<1)$. Аналогично, если на $(k-1)$-е раздражение крыса реагирует неправильно, то после обучения на k-е раздражение она с вероятностью $q_{12}=a$ реагирует правильно, а с вероятностью $q_{22}=1-a$ — неправильно $(0<a<1)$. Производится n опытов $(n>1,\ k=1, \ldots, n)$.

Какова вероятность того, что k-я реакция крысы будет правильной?

Ответ дается тем же равенством, что и в примере с частицей. В частности, если $a+b=1$ и $c=a/(a+b)$, то при каждом раздражении вероятность правильной реакции одна и та же и равна a, т. е. описываемый такой схемой процесс обучения не дает никакого эффекта.

Если $c<a/(a+b)$, то обучение дает определенный эффект. Продолжая его достаточно долго, можно произвольно приблизиться к предельной вероятности правильной реакции $a/(a+b)$. Наконец, если $c>a/(a+b)$, то обучение дает отрицательный эффект: вероятность правильной реакции уменьшается.

ЛИТЕРАТУРА

К разделу I

1. **Н. Я. Виленкин.** Рассказы о множествах. М., «Наука», 1969, 159 с.
В книге в популярной форме описываются основные понятия теории множеств.
2. **Н. Я. Виленкин.** Комбинаторика. М., «Наука», 1969, 328 с.
В книге в занимательной форме рассказывается о комбинаторных методах и задачах.
3. **Дж. Т. Калбертсон.** Математика и логика цифровых устройств. М., «Просвещение», 1965, 267 с.
В книге, в частности, описываются и используются некоторые комбинаторные методы. Книга рассчитана на школьников старших классов.
4. **Дж. Кемени, Дж. Снелл, Дж. Томпсон.** Введение в конечную математику. М., ИЛ, 1964, 486 с.
Первые главы книги посвящены элементам логики, теории множеств и комбинаторики. Книга рассчитана на студентов младших курсов.
5. **Дж. Риордан.** Введение в комбинаторный анализ. М., ИЛ, 1963, 287 с.
В книге дается систематическое описание комбинаторных методов и решается большое количество задач. Книга написана для научных работников и инженеров.

К разделу II

1. **Э. Борель.** Вероятность и достоверность. М., Физматгиз, 1961, 119 с.
В книге подробно обсуждается содержание понятия вероятности.
2. **Б. В. Гнеденко, А. Я. Хинчин.** Элементарное введение в теорию вероятностей. М., «Наука», 1964, 144 с.
Книга в простой форме и на базе конкретных примеров знакомит читателя с основными понятиями теории вероятностей.
3. **Ф. Мостеллер, Р. Рурке, Дж. Томас.** Вероятность. М., «Мир», 1969. 431 с.
Книга представляет собой элементарное введение в теорию вероятностей и статистику. В ней много интересных задач и примеров.
4. **Дж. Кемени, Дж. Синелл, Дж. Томпсон.** Введение в конечную математику. М., ИЛ, 1964, 486 с.
В последних главах книги излагаются элементы теории вероятностей и некоторые ее приложения.
5. **Дж. Т. Калбертсон.** Математика и логика цифровых устройств. М., «Просвещение», 1965, 267 с.
Одна из глав этой книги посвящена элементарной теории вероятностей.
6. **А. Кофман, Р. Фор.** Займемся исследованием операций. М., «Мир», 1966, 279 с.
При решении некоторых из рассматривающихся в книге задач используются вероятностные методы.

7. **А. М. Яглом, И. М. Яглом.** Вероятность и информация. М., «Наука», 1973, с. 511.

Книга представляет собой элементарное введение в теорию информации.

8. **А. М. Яглом, И. М. Яглом.** Неэлементарные задачи в элементарном изложении. М., Гостехиздат, 1954, 237 с.

В книге собраны, в частности, сто задач по комбинаторике и теории вероятностей с пояснениями и подробными решениями.

9. **Ф. Мостеллер.** Пятьдесят занимательных вероятностных задач с решениями. М., «Наука», 103 с.

Большинство этих задач не сложно, но есть и трудные.

10. **В. Феллер.** Введение в теорию вероятностей и ее приложения. Т. 1, 2. М., «Мир», 1967, 498, 752 с.

Книга является прекрасным учебным руководством по теории вероятностей.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

К разделу I

К разделу II

ОГЛАВЛЕНИЕ

РАЗДЕЛ II

ВЕРОЯТНОСТЬ

Лев Яковлевич Савельев

КОМБИНАТОРИКА И ВЕРОЯТНОСТЬ

Редактор *Н. Г. Рязанова*
Художественный редактор *Т. Ф. Каминина*
Художник *Н. А. Савельева*
Технический редактор *А. В. Семкова*
Корректоры *Н. Г. Примогенова, Л. Л. Тычкина*



Сдано в набор 2 августа 1974 г. Подписано в печать 7 января 1975 г. МН 02105. Формат 60×90/16. Бумага машиномелованная 26,5 печ. л., 26,4 уч.-изд. л. Тираж 16300 экз. Заказ 168. Цена 1 р. 97 к.

Издательство «Наука», Сибирское отделение. 630099, Новосибирск, 99, Советская, 18.
4-я типография издательства «Наука». 630077, Новосибирск, 77, Станиславского, 25.

0210
300 эк
18.
25.


Л. Я. САВЕЛЬЕВ • КОМБИНАТОРИКА И ВЕРОЯТНОСТЬ